

TEXTO ILUSTRADO DE

Biología Molecular e Ingeniería Genética

Conceptos, Técnicas y Aplicaciones en Ciencias de la Salud

José Luque
Ángel Herráez

TEXTO ILUSTRADO DE

**Biología Molecular
e Ingeniería Genética**

Conceptos, Técnicas y Aplicaciones
en Ciencias de la Salud

This One



GDSK-HPQ-3D7Y

TEXTO ILUSTRADO DE
Biología Molecular
e Ingeniería Genética
Conceptos, Técnicas y Aplicaciones
en Ciencias de la Salud

José Luque Cabrera
Catedrático de Universidad

Ángel Herráez Sánchez
Profesor Titular de Universidad

Departamento de Bioquímica y Biología Molecular
Universidad de Alcalá
Alcalá de Henares, Madrid



Madrid - Barcelona - Amsterdam - Boston - Filadelfia
Londres - Orlando - París - Roma - Sydney - Tokio - Toronto

Es una publicación



© MMI Elsevier España, S.A.
Génova, 17-3.º
28004 - Madrid, España

Fotocopiar es un delito (Art. 270 C.P.)

Para que existan libros es necesario el trabajo de un importante colectivo (autores, traductores, dibujantes, correctores, impresores, editores...). El principal beneficiario de este esfuerzo es el lector que aprovecha su contenido. Quien fotocopia un libro, en las circunstancias previstas por la ley, delinque y contribuye a la «no» existencia de nuevas ediciones. Además, a corto plazo, encarece el precio de los ya existentes.

Este libro está legalmente protegido por los derechos de propiedad intelectual. Cualquier uso, fuera de los límites establecidos por la legislación vigente, sin el consentimiento del editor, es ilegal. Esto se aplica en particular a la reproducción, fotocopia, traducción, grabación o cualquier otro sistema de recuperación de almacenaje de información.

ISBN: 84-8174-505-7
Depósito legal: M. 28.347-2006
Impreso en España por: Gráficas Muriel, S.A.

ADVERTENCIA

La medicina es un área en constante evolución. Aunque deben seguirse unas precauciones de seguridad estándar, a medida que aumenten nuestros conocimientos gracias a la investigación básica y clínica habrá que introducir cambios en los tratamientos y en los fármacos. En consecuencia, se recomienda a los lectores que analicen los últimos datos aportados por los fabricantes sobre cada fármaco para comprobar la dosis recomendada, la vía y duración de la administración y las contraindicaciones. Es responsabilidad ineludible del médico determinar las dosis y el tratamiento más indicado para cada paciente, en función de su experiencia y del conocimiento de cada caso concreto. Ni los editores ni los directores asumen responsabilidad alguna por los daños que pudieran generarse a personas o propiedades como consecuencia del contenido de esta obra.

EL EDITOR

A nuestros alumnos, inspiradores de este esfuerzo.

A nuestros compañeros profesores,
Montse, Cristina, José Carlos,
Pilar y Ana, por su confianza y apoyo continuos.

A nuestras familias y amigos, por su comprensión
y por el tiempo que han cedido.

Prólogo

En la Presentación de este libro, que tengo el honor de prologar, los autores afirman que pretenden reunir los conceptos actuales de la Biología Molecular y las modernas metodologías de la Ingeniería Genética con vistas a su aplicación a la Patología Molecular, humana y animal, en sus diversos aspectos: etiología, fisiopatología, diagnóstico, pronóstico, terapéutica y prevención de las enfermedades, y promoción de la salud. Y pretenden hacerlo desde una perspectiva pluridisciplinar e integradora dirigida, fundamentalmente, a los estudiantes que cursan las licenciaturas de Farmacia, Medicina, Odontología y Veterinaria, y las orientaciones biosanitarias de Biología, Bioquímica y Química. Se trata, sin duda, de una empresa tan oportuna y ambiciosa, como compleja y difícil. Pero, a la vista del resultado alcanzado, puede afirmarse que el éxito ha sido total.

La elaboración de un texto de las características mencionadas es oportuna y necesaria. En los últimos decenios, la Biología ha experimentado un proceso de progresiva "molecularización" que ha conducido a un nuevo paradigma biológico que trata de entender lo que se ha denominado "la logique du vivant" (F. Jacob, 1970) o "the molecular logic of the living state" (A. L. Lehninger, "Biochemistry", 1970); las bases moleculares subyacentes a los procesos biológicos. El reciente desciframiento del genoma de diversos organismos, y muy especialmente del genoma humano, constituye una etapa fundamental que ha de hallar su continuidad en el no menos importante estudio de los proteomas. Siguiendo los pasos de la Biología, la Medicina y la Veterinaria están experimentando un proceso de "molecularización" creciente que ha hecho emerger la que se ha llamado Patología Molecular. Hoy en día, el nivel molecular no es sólo un nivel de investigación biomédica, sino también un nivel de actuación diagnóstica, pronóstica y terapéutica. Por ello, es imprescindible preparar adecuadamente a los futuros profesionales de la salud, ya que sólo si disponemos de profesionales competentes podremos hacer frente con éxito a los retos del futuro. El ejercicio profesional de los años venideros exigirá competencia para abordar a nivel molecular los problemas de la salud. Por ello, harán falta profesionales que posean los conceptos y los elementos de lenguaje molecular necesarios para comprender las bases moleculares de la salud y la enfermedad. Profesionales capaces de adquirir constantemente la información bioquímica necesaria para mantener su competencia, y capaces de aplicar dicha información a la solución de los problemas que les conciernen; profesionales con capacidad de autoaprendizaje para poder sustituir los conceptos caducos por los nuevos, haciendo uso de todas las posibilidades que las modernas tecnologías de la información pondrán a su alcance. Para la formación básica de tales profesionales tienen gran valor los textos que resumen los nuevos elementos conceptuales de las diversas áreas de la Biología junto a las correspondientes técnicas y metodologías; ya que permiten que el estudiante adquiera una base sólida a partir de la cual podrá ampliar los conocimientos acudiendo, luego, a las monografías especializadas, a las revisiones o, incluso, a los artículos originales. Por todo ello, y por abarcar un campo hasta ahora insuficientemente cubierto, consideramos que la obra de José Luque y Ángel Herráez responde a una necesidad urgente e importante.

Ahora bien, como hemos indicado, la empresa no era nada fácil. Y no lo era por su carácter de obra básica, dirigida a estudiantes de varias licenciaturas y, por tanto, con formación diversa, y por su pretensión de seguir un tratamiento de los temas integral y multidisciplinar. Los autores pretendían integrar lo conceptual, lo tecnológico y lo aplicado, abordando lo esencial y rechazando, en áreas de la claridad, lo complejo y no bien establecido. Es evidente que estos objetivos requerían, por parte de los autores, tanto un conocimiento profundo y actualizado de los diversos temas tratados como un conocimiento amplio y completo de los contenidos de la Bioquímica y Biología Molecular, y un conocimiento de los contenidos formativos de las diversas licenciaturas; a la vez que exigían una gran capacidad didáctica. Sin duda, la prolongada e intensa dedicación de los autores a la docencia universitaria ha sido un factor decisivo para culminar con éxito la empresa acometida. El método de trabajo seguido para la elaboración del texto, desde la recogida de la información inicial a la revisión final de cada capítulo, teniendo en cuenta las observaciones y sugerencias de los profesores, especialistas, doctorandos e incluso estudiantes de licenciatura que lo han analizado, ha contribuido, también sin duda, a la excelencia de la labor realizada. Excelencia acrecentada por el acertado tratamiento visual que se ha hecho de la información presentada.

En resumen, el "Texto Ilustrado de Biología Molecular e Ingeniería Genética" de los Profesores José Luque y Ángel Herráez es una obra ambiciosa, muy bien realizada y que ha de ser de gran utilidad para la formación tanto de la amplia gama de estudiantes a quienes va dirigida como de los profesionales de la salud que deseen actualizar sus conocimientos en uno de los campos más importantes de la actual biomedicina. Es por todo ello, que ha sido para mí un honor y una satisfacción redactar estas líneas como Prólogo.

PROF. JOSÉ CARRERAS BARNES
UNIVERSIDAD DE BARCELONA
DICIEMBRE 2000

Presentación

OBJETIVO

Los avances y aplicaciones de la investigación bioquímica y genética ocupan hoy día una posición privilegiada en el campo de las ciencias de la salud en general y en la Farmacia y Biomedicina en particular. Se vislumbra en los albores del siglo XXI una tendencia hacia el enfoque molecular de la enfermedad, como referente último en la interpretación de los procesos patológicos y como horizonte de nuevas estrategias diagnósticas y terapéuticas basadas en dichos avances.

Para dar respuesta a esta situación se requiere una adecuada preparación universitaria, lo que justifica el desarrollo del presente proyecto: un libro de texto que facilite la adquisición de una formación básica en el área de **Biología Molecular**, un enfoque estructural y tridimensional del material genético y de los productos génicos, así como en la **Ingeniería Genética**, en un sentido amplio la modificación de dicho material en beneficio de la salud o de la técnica. Por su gran significado y su potencial futuro, estas disciplinas ya se enseñan, con mayor o menor profundidad, en las licenciaturas con responsabilidad directa o indirecta en la sanidad (Farmacia, Medicina, Odontología y Veterinaria; Biología, Bioquímica y Química) y en determinadas diplomaturas (Enfermería y Fisioterapia). Con el fin de contribuir a esta formación, se propone un enfoque adaptado a la realidad, asequible a todo este alumnado, que pueda facilitar el proceso de enseñanza y aprendizaje de tan importantes y novedosas materias en un área de conocimiento en continua expansión.

PLANTEAMIENTO

Para dar cumplimiento a este objetivo, el libro se plantea bajo una triple vertiente:

Conceptos

Dada la variedad y particularidades del alumnado al que va dirigida la obra, se ha puesto un énfasis especial en los fundamentos, comunes para todos. Se discuten con sencillez, aunque con el rigor científico exigido hoy día, las bases teóricas imprescindibles para comprender la materia y para desarrollar y aplicar científicamente la tecnología derivada de estos conceptos. Se pretende un libro válido para conocer de forma adecuada y suficiente la Biología Molecular e Ingeniería Genética sin tener que acudir a una excesiva, y a veces innecesaria, especialización.

Aspectos técnicos

La metodología derivada de la Biología Molecular, y en especial de la Ingeniería Genética (inicialmente limitada a la tecnología del DNA recombinante), se ha ido extendiendo con el tiempo a fines muy diversos (aislamiento, clonación, hibridación, secuenciación, etc.), para dar lugar a una *Biotecnología Molecular*, término que en cierto modo podría utilizarse como alternativa a *Ingeniería Genética* en el título del libro. De ahí que, bajo un punto de vista formativo, nuestro segundo objetivo consista en estudiar los principios de los métodos y técnicas más comunes, aquéllos de mayor impacto y utilidad en la investigación básica y en el laboratorio clínico. La descripción y aplicación concretas de las variantes tecnológicas, metodológicas o instrumentales, cada vez más diversas y en modificación constante, debe ser relegada para los periodos de doctorado o de especialización profesional en áreas del laboratorio clínico, entre otras, principalmente en Análisis Clínicos o en Bioquímica Clínica y Patología Molecular.

Aplicaciones biosanitarias

La Biología Molecular e Ingeniería Genética ha contribuido a resolver problemas de carácter industrial (eliminación de xenobióticos o residuos), agrícola y ganadero (mejora genética y alimentos transgénicos). Las aplicaciones en Farmacia, Medicina y Veterinaria, que se centran en la lucha contra la enfermedad humana y animal (etiología, patogenia, diagnóstico, tratamiento y prevención) y en la síntesis de moléculas con fines terapéuticos, ya han rendido sus primeros frutos. Este extraordinario avance ha dado lugar a una serie de términos que, en general, incluyen el prefijo *bio-* o el adjetivo *molecular*. Como representativo de todos ellos se acepta hoy el término *Biomedicina* o *Medicina Molecular*, considerado de una forma amplia como una "visión molecular" de todos los aspectos relacionados con la enfermedad.

CARACTERÍSTICAS

Tratamiento integral y multidisciplinar

Se ha buscado integrar las técnicas y aplicaciones con los conceptos, abordando lo esencial y rechazando, salvo excepciones, lo complejo o no bien establecido, relacionando cada nueva información con lo discutido en otros temas y con los ámbitos bioquímico, genético, químico, fisiológico, analítico, etc. Así lo sugiere la insuficiente o desproporcionada

da preparación recibida en ocasiones por el alumno en materias básicas previas, la inclusión de la Biología Molecular e Ingeniería Genética dentro de una disciplina de Bioquímica General, su impartición en programas variados, en cursos distintos, con un número limitado de créditos docentes, etc. Se ha intentado proporcionar al libro un cierto carácter autónomo, adaptable al ritmo, nivel y situación de cada estudio particular. Ello no obsta, muy al contrario, para que deba seguir recomendándose la consulta de la bibliografía específica y, sobre todo, de los excelentes libros de texto disponibles hoy día, con un abanico y tratamiento de los temas diferentes al nuestro.

Diseño visual

Como aspecto más significativo, y para facilitar la característica anterior, el libro se apoya en un tratamiento visual, original, con profusión de esquemas y figuras predominando sobre un texto lo más claro y sencillo posible. Como consecuencia, las ilustraciones deben considerarse parte integrante del texto por lo que, en contraste con lo habitual, no son citadas en éste ni tienen un pie o título descriptivo. El lector debe acostumbrarse a examinar la figura cuando llegue a ella, pues la posición de ésta en la página es exactamente aquélla que le corresponde en la lectura (salvo contadas excepciones debidas al tamaño de la figura). Por otra parte, se hace un uso extensivo del color, no sólo para aumentar el atractivo de las ilustraciones, sino para transmitir una información; por ejemplo, el objeto dibujado y el texto relativo a él comparten color. Los dibujos son originales en su concepción y pretenden la interpretación pedagógica de descripciones que por escrito resultan a veces complicadas o tediosas, o de los dibujos y esquemas de otros autores. Se persigue un carácter intuitivo, que facilite la comprensión y motive al alumno al estudio detenido de la materia, al tiempo que enriquece la exposición. Sería deseable que este esfuerzo pudiera servir para vencer la resistencia actual del alumnado a la consulta y contraste de las fuentes bibliográficas, para convencerlo de lo antiuniversitario del uso indiscriminado de los apuntes, propios o ajenos, como única fuente de información. Como estímulo adicional a este enfoque, la Editorial también ha tratado de lograr un precio universitario.

Terminología

Se han limitado las palabras inglesas y sus traducciones excesivamente literales, empleando los términos castellanos existentes. A pesar de ello, la definición de algunos se acompaña de la equivalencia inglesa, cuando ésta es de uso muy extendido en las fuentes bibliográficas y, especialmente, cuando permite comprender los acrónimos. A este respecto, hemos decidido utilizar exclusivamente los acrónimos ingleses, dado su uso extensivo en toda la bibliografía y en el lenguaje científico hablado; nos parece, por ejemplo, que tiene escasa utilidad que el alumno conozca la 'RCP', cuando en todos los entornos profesionales en que se pueda mover en el futuro se encontrará con la 'PCR'. Como consecuencia más llamativa de este criterio, los ácidos nucleicos no se llaman ADN y ARN, sino DNA y RNA.

Enfoque a organismos eucariotas

Nuestro tratamiento de la materia tiende a recoger el máximo de información posible, aún dispersa y escasa en otras fuentes, relativa a mamíferos en general y a humanos en particular. Se entiende, por ejemplo, que debe ponerse más énfasis en el cromosoma humano y sus alteraciones que en el mecanismo de acción de las DNA polimerasas de *Escherichia coli*. Obviamente, no se renuncia a la gran riqueza de conocimientos alcanzada en procariotas y virus, sobre todo cuando esta información es esencial para comprender procesos teóricos o de interés analítico en eucariotas.

Orientación

El contenido se desarrolla pensando en las orientaciones biosanitarias que posiblemente más requieren del conocimiento de los conceptos y técnicas de la Biología Molecular e Ingeniería Genética, aunque, dados su amplitud y su futuro, en el presente libro sólo pueden apuntarse las bases mínimas necesarias para su comprensión. Por un lado, el **diagnóstico molecular**, ya de uso casi rutinario en los modernos laboratorios clínicos para, por ejemplo, la prevención prenatal, la identificación de individuos con fines forenses, etc. Debe, no obstante, señalarse que no se trata de un libro de Patología Molecular, sino un texto básico sobre los conocimientos necesarios para abordar procesos patológicos. Por otra parte, la **terapia molecular**, o tratamiento de futuro de la enfermedad basado en los avances sobre modificación genética de las células (terapia génica), el conocimiento del genoma humano (genómica) y, más a largo plazo, su relación con la funcionalidad de las proteínas (proteómica).

Desarrollo del libro

El método de trabajo se ha basado en una total corresponsabilidad y coordinación de ambos autores a lo largo de su construcción (en su más estricto sentido), en varias etapas durante los casi cuatro años de su gestación. Se ha partido de toda la información recogida previamente, con fines docentes, de innumerables fuentes (textos, libros especializados, revisiones, artículos originales, notas personales, etc.). Cada tema, una vez redactada su primera versión, en ocasiones a partir de trabajos en grupo con estudiantes, se readaptó incorporando las figuras originales elaboradas para la ocasión e incluyendo en ellas toda la información posible a expensas del texto. En una segunda etapa, dicho trabajo se ha visto enriquecido, pedagógicamente y científicamente, por la labor de revisión, observaciones minuciosas y críticas rigurosas realizadas por antiguos alumnos, doctorandos, compañeros profesores, investigadores y profesionales. Por ello, con nuestro profundo agradecimiento, los nombres de estas personas aparecen citados en la obra.

Revisión y asesoramiento

Revisores

Nos es grato hacer mención del esfuerzo generoso realizado por las personas que han revisado temas concretos del libro, en alguno de los momentos de su elaboración. A pesar de esta colaboración, la responsabilidad de los posibles errores es sólo nuestra.

Miguel Aguilar Urbano

Profesor Asociado,
Dpto. Bioquímica y Biología Molecular
Universidad de Córdoba
(Temas 12 y 15)

Francisco Javier Álvarez Pachón

Doctor en Farmacia, Industria farmacéutica
Alcalá de Henares
(Tema 15)

Eduardo Arilla Ferreiro

Catedrático, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 1)

M^a Dolores de Arriaga Giner

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de León
(Temas 2, 3 y 16)

Milagros Ballesteros González

Profesora Titular,
Dpto. Química Analítica e Ingeniería Química
Universidad de Alcalá
(Tema 11)

M^a Carmen Boyano Adán

Profesora Titular Interina,
Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Temas 12 y 16)

Xosé R. Bustelo

Científico Titular, Centro de Investigación del Cáncer
Universidad de Salamanca – CSIC
(Tema 30)

Fernando Climent Romeo

Catedrático, Dpto. Ciencias Fisiológicas I
Universidad de Barcelona
(Tema 13)

M^a Dolores Delgado Villar

Profesora Titular, Dpto. Biología Molecular
Universidad de Cantabria
(Temas 8 y 9)

José Carlos Díez Ballesteros

Profesor Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 7)

Isabel Fabregat Romero

Profesora Titular,
Dpto. Bioquímica y Biología Molecular II
Universidad Complutense de Madrid
(Temas 20, 21 y 26)

Margarita Fernández de Castro

Profesora Titular,
Dpto. Bioquímica y Biología Molecular II
Universidad Complutense de Madrid
(Temas 20, 21 y 26)

M^a Dolores Fernández Moreno

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 4)

Esther Ferrer Cebrián

Profesora Titular, Dpto. Biología Celular y Genética
Universidad de Alcalá
(Temas 17 y 18)

Araceli Fominaya Yagüe

Profesora Titular, Dpto. Biología Celular y Genética
Universidad de Alcalá
(Temas 17 y 18)

Luis Franco Vera

Catedrático, Dpto. Bioquímica y Biología Molecular
Universidad de Valencia
(Tema 19)

Xavier Fuentes Arderiu

Servicio de Bioquímica Clínica
Ciudad Sanitaria y Universitaria de Bellvitge,
Hospitalet de Llobregat
(Temas 28 y 29)

Juan Pedro García Ballesta

Profesor de Investigación,
Centro de Biología Molecular "Severo Ochoa"
Universidad Autónoma de Madrid – CSIC
(Tema 23)

Ana I. García Pérez

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 3)

Marina Garin Ferreira

Investigadora posdoctoral, Dpt. Haematology,
Imperial College School of Medicine
Hammersmith Hospital, Londres
(Tema 22)

Luis A. González Guijarro

Profesor Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 5)

José M. González de Buitrago

Catedrático, Dpto. Bioquímica y Biología Molecular
y Jefe de Sección de Bioquímica, Hospital Universitario
Universidad de Salamanca
(Temas 28 y 29)

Pilar Iniesta Cano

Profesora Titular,
Dpto. Bioquímica y Biología Molecular II
Universidad Complutense de Madrid
(Temas 20, 21 y 26)

Antonio Jiménez Ruiz

Profesor Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 6)

M^a Carmen Lastra González

Profesora Titular,
Dpto. Química Analítica e Ingeniería Química
Universidad de Alcalá
(Tema 11)

Javier León Serrano

Catedrático, Dpto. Biología Molecular
Universidad de Cantabria
(Temas 8 y 9)

Juan López Barea

Catedrático, Dpto. Bioquímica y Biología Molecular
Universidad de Córdoba
(Temas 5 y 6)

José M^a Macarulla Greoles

Catedrático, Dpto. Bioquímica y Biología Molecular
Universidad del País Vasco
(Tema 12)

Mercedes Martín Marquín

Profesora Titular, Dpto. Biología Vegetal
Universidad de Alcalá
(Tema 21)

Honorina Martínez Blanco

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de León
(Temas 2, 3 y 16)

M^a Begoña Martínez Jarreta

Profesora Titular, Laboratorio de Genética Forense,
Dpto. Anatomía Patológica, Medicina Legal y Forense
y Toxicología
Universidad de Zaragoza
(Tema 8)

Alberto Martínez Serrano

Profesor Titular,
Centro de Biología Molecular "Severo Ochoa"
Universidad Autónoma de Madrid - CSIC
(Tema 5)

Miguel Ángel Medina Torres

Profesor Titular, Dpto. Biología Molecular y Bioquímica
Universidad de Málaga
(Tema 30)

Jesús Mendieta Gómez

Investigador contratado, Centro de Biología Molecular
"Severo Ochoa" y Dpto. Farmacología
CSIC y Universidad de Alcalá
(Temas 6 y 14)

Julio Montoya Villaroya

Profesor Titular,
Dpto. Bioquímica y Biología Molecular y Celular
Universidad de Zaragoza
(Tema 19)

Jorge Monserrat Sanz

Ldo. en Biología y Técnico de Citometría,
Dpto. Medicina
Universidad de Alcalá
(Tema 11)

Felipe Moreno Herrero

Jefe de la Unidad de Genética Molecular
Hospital Ramón y Cajal, Madrid
(Tema 10)

M^a Teresa Muño Blanco

Profesora Titular,
Dpto. Bioquímica y Biología Molecular y Celular
Universidad de Zaragoza
(Tema 11)

Rafael Oliva Virgili

Servicio de Genética
Hospital Clínico, Barcelona
(Temas 17 y 18)

Fidel Ortega Ortiz de Apodaca

Profesor Titular,
Dpto. Química Analítica e Ingeniería Química
Universidad de Alcalá
(Tema 11)

Acisclo Pérez Martos

Profesor Titular,
Dpto. Bioquímica y Biología Molecular y Celular
Universidad de Zaragoza
(Temas 1, 10 y 18)

Luis del Peso Ovale

Investigador del Servicio de Inmunología
Hospital Universitario de la Princesa, Madrid
(Tema 13)

Pedro Piedras Montilla

Investigador Contratado de Reincorporación,
Dpto. Bioquímica y Biología Molecular
Universidad de Córdoba
(Temas 12 y 15)

Manuel Pineda Priego

Profesor Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Córdoba
(Temas 12 y 15)

Montserrat Pinilla Barrau

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 4)

Alfredo Prieto Martín

Profesor Asociado, Dpto. Medicina
Universidad de Alcalá
(Tema 11)

Lilian Puebla Jiménez

Profesora Titular Interina,
Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 16)

Carmen Pueyo de la Cuesta

Catedrática, Dpto. Bioquímica y Biología Molecular
Universidad de Córdoba
(Tema 25)

Juan A. Pulido Mora

Doctor en Ciencias Biológicas y
Supervisor de Instalaciones Radiactivas
(Temas 16 y 24)

Ángel Reglero Chillón

Profesor Titular, Dpto. Bioquímica y Biología Molecular
Universidad de León
(Temas 2, 3 y 16)

Eduardo Reyes Martín

Profesor Asociado, Dpto. Medicina
Universidad de Alcalá
(Tema 11)

Nieves Rodríguez Henche

Profesora Titular Interina,
Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 16)

José Carlos Rodríguez Rey

Profesor Titular, Dpto. Biología Molecular, Unidad
Asociada al Centro de Investigaciones Biológicas
Universidad de Cantabria – CSIC
(Tema 15)

I. Dolores Román Curto

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 3)

Bartolomé Sabater García

Catedrático, Dpto. Biología Vegetal
Universidad de Alcalá
(Tema 21)

Francisca Sánchez Jiménez

Profesora Titular, Dpto. Biología Molecular y Bioquímica
Universidad de Málaga
(Tema 30)

José Sánchez-Prieto Borja

Profesor Titular,
Dpto. Bioquímica y Biología Molecular IV
Universidad Complutense de Madrid
(Temas 20, 21 y 26)

Pilar Sancho López

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Tema 2)

Juan Saus Más

Investigador Jefe del Dpto. Patología Molecular
Fundación Valenciana de Investigaciones Biomédicas,
Instituto de Investigaciones Citológicas, Valencia
(Tema 13)

Joaquín Soler Medina

Profesor Titular, Dpto. Bioquímica y Biología Molecular
Universidad de León
(Temas 2, 3 y 16)

M^a Cristina Tejedor Gilmartín

Profesora Titular, Dpto. Bioquímica y Biología Molecular
Universidad de Alcalá
(Temas 5 y 22)

Miguel Ángel Trueba Conde

Profesor Titular, Dpto. Bioquímica y Biología Molecular
Universidad del País Vasco
(Tema 12)

Alumnos de licenciatura y de doctorado

Varios temas se iniciaron como fruto del trabajo en grupo con alumnos, a partir del curso 1997/98. Algunos de ellos han revisado posteriormente con especial dedicación temas concretos.

Gema Alonso Poves

Licenciada

Paloma Calleja Lamparero

Licenciada; industria farmacéutica

Susana Cantero Blanco

Licenciada; residente hospitalaria

Diana de Córdoba del Campo

Licenciada; industria farmacéutica
(Tema 14)

Mónica Farto Paredes

Becaria predoctoral

Raquel Fuentes Irigoyen

Licenciada; residente hospitalaria

Inmaculada Fuentes Rus

Licenciada

Raquel García Dávila

Licenciada; industria farmacéutica
(Tema 14)

Celia García Moreno

Licenciada; industria farmacéutica

Carlos David Gea Méndez

Licenciado

Francisco Javier González Cappa

Licenciado
(Temas 8 y 9)

Beatriz González Díaz

Licenciada

Teresa Gonzalo Lázaro

Ex-alumna

Marta Herranz Berrón

Licenciada

Maria Paz Jiménez Hernández

Licenciada

José Antonio Jordán Hernández

Doctor; profesor de enseñanza secundaria

Carmen Lizano Vaquero

Doctora; industria farmacéutica

Luis Alfredo Lotero Martín

Licenciado; industria farmacéutica

Luisa Lucas Parras

Investigadora posdoctoral

Antonio José Martínez Osorio

Licenciado; industria farmacéutica

Paula Mascías Núñez

Becaria predoctoral

(Tema 11)

Elena Moreno Centeno

Ex-alumna

Juan Carlos Murciano Fernández

Investigador posdoctoral

M^a Inmaculada Navarro Pérez

Licenciada

Gemma Olmos Centenera

Becaria predoctoral

Sara Ortiz Garrudo

Licenciada; industria farmacéutica

M^a Teresa Pérez Maroto

Doctora; laboratorio clínico

Francisco Javier Rivero Vilches

Becario predoctoral

(Temas 6, 11 y 12)

Helena Rodero Torres

Licenciada; industria farmacéutica

Fátima Rodríguez Barrios

Becaria predoctoral

(Temas 11 y 18)

Alicia Rodríguez-Bobada Jiménez

Ex-alumna

José Alberto Rojo López

Licenciado

(Tema 25)

Sara del Rosal Monedero

Licenciada

Ana María Sánchez de Abajo

Ex-alumna

Berta Sánchez de Juan

Ex-alumna

Rosa Sánchez Sánchez

Licenciada; industria farmacéutica

Silvia Sanz Moratilla

Doctora; farmacéutica titular

Omar Sharaf Ordóñez

Ex-alumno

M^a Ángeles Villar Núñez

Becaria predoctoral

(Temas 6, 11 y 12)

Isabel Zardain Rodríguez

Ex-alumna

Otros agradecimientos

Tanto al inicio como durante el desarrollo del libro, diversas personas han mostrado su apoyo y participado con comentarios, críticas, sugerencias al temario y al planteamiento global, por lo que también queremos expresarles nuestra gratitud:

Jesús Ávila de Grado (Madrid), **José A. Cabezas Fernández del Campo** (Salamanca), **José Carreras Barnes** (Barcelona), **Antonio Casas Moreno** (Madrid), **José A. Cebrián Pérez** (Zaragoza), **Cristina Delgado Montes** (Londres), **Eduardo García Peregrín** (Granada), **Ángel Giménez Solves** (Madrid), **Nicolás Jouve de la Barreda** (Alcalá de Henares), **Manuel J. López Pérez** (Zaragoza), **José Manuel Luque Pinilla** (Madrid), **Ignacio Núñez de Castro** (Málaga), **Manuel Ortega Mata** (Madrid), **Jacinto Pinilla Barrau** (Zaragoza), **Pedro Roncalés Rabinal** (Zaragoza), **Vicente Rubio Zamora** (Valencia) y **Enrique Villar Ledesma** (Salamanca).

Agradecemos igualmente a **Elsevier España, S.A.**, en la persona de su directora editorial, **María José Hernández Gómez**, su confianza en este proyecto y su empeño en que el libro haya salido adelante con todas sus peculiaridades técnicas.

Índice

PRELIMINARES

Prólogo *IX*

Presentación *XI*

Revisión y asesoramiento *XIII*

I. ESTRUCTURA Y FUNCIÓN DEL MATERIAL GENÉTICO

Tema 1:

Introducción y aspectos generales en eucariotas 1

- Transmisión de la información genética 1
- El DNA como material genético: Antecedentes 2
- Características generales del genoma 5

Tema 2:

Componentes fundamentales de los ácidos nucleicos 9

- Componente ácido: Fosfatos 9
- Componente neutro: Azúcares 11
- Componente básico: Bases nitrogenadas 12

Tema 3:

Nucleósidos y nucleótidos 19

- Estructura de nucleósidos 19
- Estructura de nucleótidos 23

Tema 4:

Estructura primaria de ácidos nucleicos 31

- Aspectos generales 31
- Estructura primaria 32
- Formas de representación lineal 32
- Dos tipos de ácidos nucleicos según su composición 33
- Propiedades fisicoquímicas de los ácidos nucleicos 34

Tema 5:

Estructura secundaria del B-DNA 37

- Proporción de bases nitrogenadas: Reglas de Chargaff 38
- Modelo de Watson y Crick: forma B del DNA 39

Tema 6:

Variaciones en la estructura del DNA 51

- Variantes en doble hebra: formas A y Z 52
- Variantes locales de la estructura secundaria del B-DNA 56
- Motivos estructurales responsables de la unión del DNA con proteínas 60

Tema 7:

Estructuras de orden superior de DNA y RNA 65

- Superenrollamiento del DNA 66
- Estructura de los RNAs 69
- Los ribosomas 79

Tema 8:

Condensación del DNA y cromosomas 83

- Condensación del DNA en eucariotas 83
- Estudio del cromosoma metafásico 88
- Dotación genética en eucariotas 93

Tema 9:

Ciclo celular 97

- Introducción 97
- La cromatina en la interfase 98
- División de células somáticas por mitosis 99
- División de células germinales primarias por meiosis 99
- Unión de los gametos haploides durante la fecundación 106
- Resumen integrado 106

Tema 10:

Organización del genoma eucariótico 107

- Complejidad del genoma eucariótico 107
- DNA de copia única, simple o no repetitivo 108
- DNA repetitivo 109
- Estudio experimental de la complejidad en el genoma de eucariotas 115

Tema 11:

Preparación de muestras, extracción y análisis de ácidos nucleicos 117

- Obtención y preparación preliminar de las muestras 119
- Disociación de la muestra tisular 122
- Separación de células 122
- Caracterización celular y medidas de viabilidad 129
- Lisis de las células 132
- Preparación de fracciones subcelulares 133
- Tratamientos adicionales o complementarios 133
- Extracción de ácidos nucleicos por solubilidad en fases inmiscibles 134
- Purificación de ácidos nucleicos por precipitación salina diferencial 135
- Recogida y conservación de las muestras 135
- Procedimientos alternativos: extracción o análisis directos del DNA 135
- Fraccionamiento de ácidos nucleicos 136

II. TRANSMISIÓN DE LA INFORMACIÓN GENÉTICA Y TECNOLOGÍAS RELACIONADAS

Tema 12:**Replicación del DNA** 141

- Características generales de la replicación 141
- Enzimología de la replicación 145
- Etapas en el proceso de replicación 150
- Replicación mitocondrial 160

Tema 13:**Hibridación de ácidos nucleicos: fundamento y métodos de ensayo** 163

- Introducción: Desnaturalización y renaturalización del DNA 163
- Análisis molecular de la hibridación de ácidos nucleicos 167
- Métodos de ensayos de hibridación 169

Tema 14**Preparación y marcaje de sondas para hibridación** 175

- Preparación de sondas genéticas 175
- Marcaje de las sondas 180

Tema 15**Clonación acelular: Reacción en cadena de la polimerasa (PCR)** 187

- Introducción general a la clonación 187
- Amplificación in vitro del DNA: Reacción en cadena de la polimerasa o PCR 188

Tema 16:**Clonación celular: Tecnología del DNA recombinante** 197

- Enzimas de restricción 197
- Clonación celular de moléculas de DNA 203
- Genotecas 215

Tema 17**Genómica. Obtención de los mapas genético y físico del genoma** 217

- Introducción 217
- Mapas genéticos 218
- Mapas físicos 221

Tema 18**Secuenciación del genoma y Proyectos Genoma** 231

- Introducción 231
- Método químico de secuenciación del DNA 231
- Método enzimático de secuenciación del DNA 235
- Secuenciación del RNA 238
- Proyecto Genoma Humano (PGH): Características esenciales 239
- Otros proyectos de secuenciación del genoma humano 240
- Genómica comparativa 241

III. EXPRESIÓN GÉNICA

Tema 19:**Transcripción** 243

- Introducción: conceptos generales 243
- Enzimología de la transcripción: Mecanismo de la reacción RNA polimerasa 246
- Transcripción en eucariotas: diferencias con procariontes 250
- Etapas en el proceso de transcripción 252
- Transcripción del genoma mitocondrial 257
- Inhibidores de la transcripción 258

Tema 20:**Control de la expresión génica: pretranscripcional y transcripcional** 261

- Introducción general a la regulación de la expresión génica 261
- Control pretranscripcional 263
- Regulación genética de la transcripción 264
- Regulación epigenética 271

Tema 21:**Maduración del RNA o procesamiento postranscripcional** 275

- Introducción 275
- Características diferenciales de la maduración 276
- Procesamiento del RNA mensajero 277
- Procesamiento de los RNAs ribosómico y de transferencia 282
- Maduración del RNA mitocondrial 284
- Regulación postranscripcional y pretraduccional de la expresión génica 284

Tema 22:**El código genético** 287

- Antecedentes y propiedades generales del código 287
- Asignación de codones a aminoácidos concretos 289
- Modelos de representación 293
- Características específicas 294

Tema 23:**Síntesis de proteínas: Traducción** 299

- Características de la traducción 300
- Fase previa: activación de los aminoácidos en forma de aminoacil-tRNAs 302
- Fase 1: iniciación 307
- Fase 2: elongación o alargamiento de la cadena peptídica 311
- Fase 3: terminación 314
- Energética de la síntesis de proteínas 315
- Inhibidores de la traducción 316
- Regulación de la síntesis proteica: Niveles de control 318

Tema 24:**Modificaciones postraduccionales** 321

- Introducción 322
- Tráfico o destino de las proteínas 322
- Maduración o procesamiento del polipéptido naciente 326
- Plegamiento de proteínas 337
- Degradación de las proteínas 340

IV. ASPECTOS APLICADOS

Tema 25:**Bases moleculares de la mutación y la reparación del DNA** 343

- Concepto de mutación 344
- Clasificación de las mutaciones 344
- Tipo de célula que sufre la mutación 344
- Mutaciones a pequeña escala o puntuales 346
- Causas y mecanismos básicos de las mutaciones 352
- Reparación del DNA 356
- Enfermedades humanas asociadas a la reparación 363

Tema 26:**Diversidad del genoma: polimorfismos** 365

- Introducción 365
- Diversidad genética dentro de una especie 366
- Mecanismos implicados en la variabilidad genética 366
- Consecuencias funcionales del polimorfismo 368

Tema 27:**Análisis de genes. Detección y aplicaciones de los polimorfismos** 379

- Introducción 379
- Análisis de genes 380
- Detección del polimorfismo de secuencia de DNA 381
- Aplicaciones de los polimorfismos de DNA 384

Tema 28:**Enfermedades moleculares: enfermedades monogénicas** 389

- Concepto y clasificación 389
- Enfermedades exógenas 390
- Enfermedades genéticas 390
- Enfermedades poligénicas, multifactoriales, mixtas o complejas 399

Tema 29:**Enfermedades cromosómicas o citogenéticas** 401

- Introducción 401
- Clasificación y nomenclatura 402
- Cromosomopatías numéricas 403
- Cromosomopatías estructurales 406

Tema 30:**Bases moleculares del cáncer** 411

- Introducción 412
- El cáncer como enfermedad genética 412
- Etapas características en el desarrollo del cáncer 413
- Mecanismos de transformación de célula normal en tumoral 414
- Control de la proliferación por señales extracelulares 420
- Regulación del ciclo celular 422
- Apoptosis, muerte celular programada o suicidio celular 431
- Herencia del cáncer 434
- Marcadores tumorales 434

APÉNDICES

Referencias históricas 441**Bibliografía** 447

ÍNDICE ANALÍTICO

Índice analítico 449

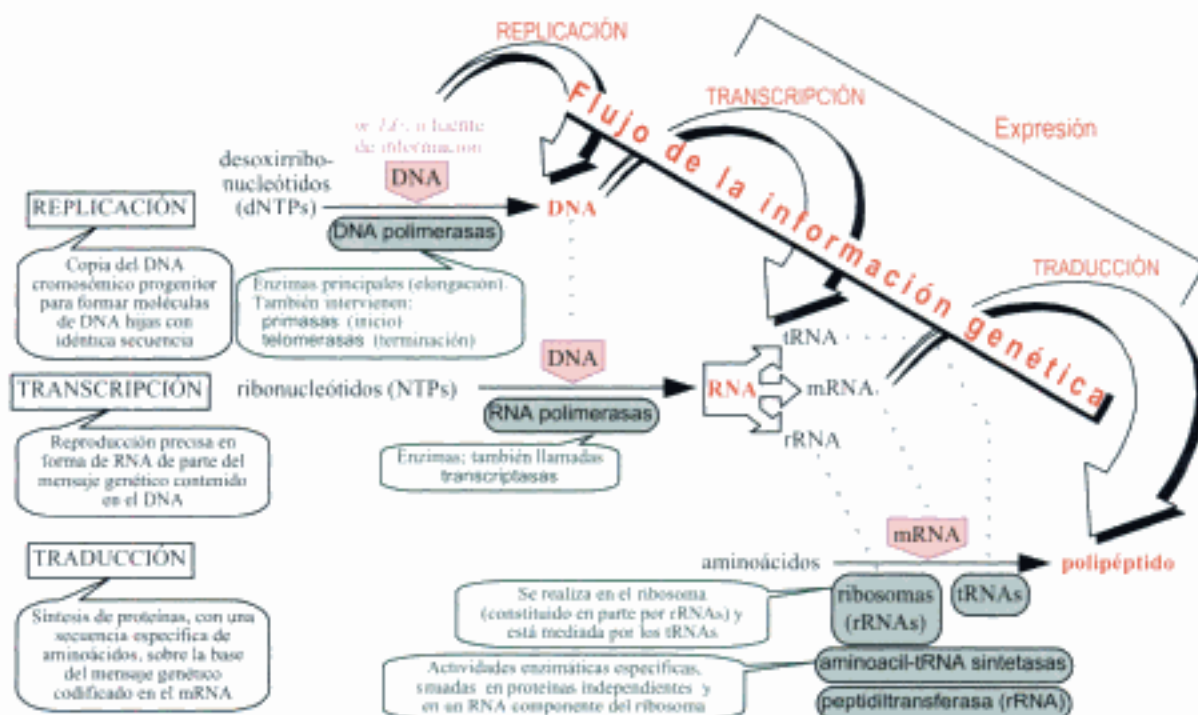
Tema 1

Introducción y aspectos generales en eucariotas

1.1	Transmisión de la información genética	1
1.2	El DNA como material genético: pasado, presente y futuro	2
1.3	Características generales del genoma	5
1.3.1	Genoma de procariotas y eucariotas	5
1.3.2	Magnitud del genoma nuclear de eucariotas	6
1.3.3	Magnitud y características del genoma mitocondrial	6

A modo de introducción general al presente texto (y, en definitiva, a nuestro enfoque de la Biología Molecular e Ingeniería Genética), es conveniente describir sucintamente los tres aspectos que mejor ponen de manifiesto el papel de los ácidos nucleicos en la célula y en especial del DNA como portador de la información genética: Por un lado, la descripción global de las rutas de transmisión de la información genética; en segundo lugar, los antecedentes experimentales que demostraron el importante papel de los ácidos nucleicos; finalmente, las características generales del genoma de eucariotas, como punto de partida para el estudio posterior en profundidad de los conceptos teóricos, las tecnologías y las aplicaciones de esta materia en las ciencias de la salud.

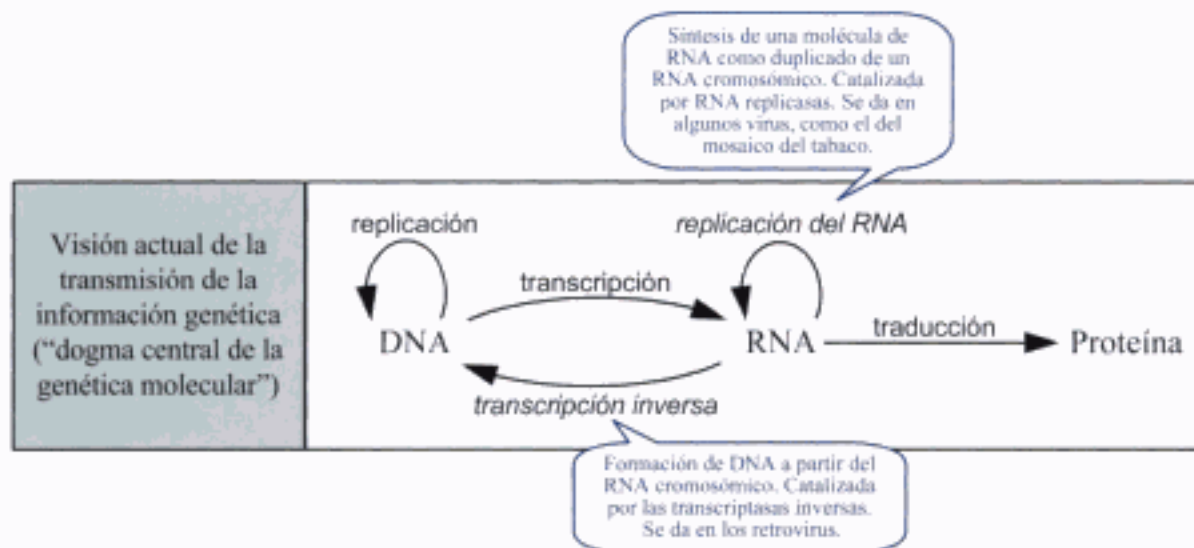
1.1 TRANSMISIÓN DE LA INFORMACIÓN GENÉTICA



Francis Crick introdujo el término "Dogma central de la genética molecular" para describir, incluso antes de conocerse su importancia real en los seres vivos, el flujo de la información genética y la utilización celular de dicha información. En concreto, la transmisión de la información contenida en la secuencia del DNA a las células y organismos descendientes mediante la replicación, y su expresión en biomoléculas funcionales mediante los procesos de transcripción y traducción. Comprende, pues, el conjunto de relaciones generales existentes entre el DNA, el RNA y las proteínas.

Debe resaltarse que la representación abreviada **DNA** → **RNA** → **Proteína** no indica la transformación de una molécula en otra, tal como se indica normalmente en las rutas metabólicas, sino que cada elemento en la secuencia aporta la información necesaria para la síntesis del siguiente. En este sentido, es importante el concepto de *plantilla* o *molde* molecular: aunque cada paso está catalizado por una enzima, los sustratos (nucleótidos, aminoácidos) no son seleccionados por ella, sino que vienen especificados por la intervención del molde, en concreto por su secuencia de nucleótidos.

El progreso en el conocimiento de los procesos de transmisión de la información genética obligó a la ampliación de este esquema básico original para acomodar otros procesos que también ocurren, al menos en algunos organismos particulares, como es el caso de algunos virus:



1.2 EL DNA COMO MATERIAL GENÉTICO: PASADO, PRESENTE Y FUTURO

El conocimiento de la estructura de los ácidos nucleicos se inicia en el siglo XIX con el aislamiento en el núcleo celular de una **nucleína** formada por una parte ácida (hoy, DNA) y una parte básica (proteína), así como con el estudio parcial de las propiedades de la nucleína y de su relación con la herencia celular. Sin embargo, la estructura sólo se conoció a mediados del siglo XX, al mostrarse que el DNA era el componente cromosómico depositario de la información genética. Cabe destacar en este sentido, como antecedentes clave, los experimentos de Oswald T. Avery, Colin MacLeod y Maclyn McCarty, en 1944, y de Alfred D. Hershey y Martha Chase, en 1952, que constituyen un hito en el conocimiento del material genético a escala molecular.

Tras estas aportaciones iniciales, el verdadero inicio de la Biología Molecular moderna lo constituyó la propuesta de estructura en doble hélice para el DNA, formulada, en 1953, por James D. Watson y Francis Crick. A partir de este momento, casi todo el esfuerzo realizado se centra en el estudio del DNA genómico, desde el punto de vista estructural y funcional, en procariotas y eucariotas, con una atención especial a mamíferos en general y humanos en particular. Los avances logrados, tanto teóricos como técnicos y aplicados, han sido enormes. A ello ha contribuido la aparición de la ingeniería genética, como principal área derivada de la biología molecular. Queda pendiente para un futuro próximo la conclusión del Proyecto Genoma Humano (genómica funcional), el análisis de las funciones de cada gen (genómica funcional) y la disección de la estructura y función de las proteínas (proteómica). Como expresión de estos nuevos conocimientos, ha surgido una gran variedad de campos de actividad en las Ciencias de la Salud (Medicina, Farmacia y Veterinaria), tales como anticuerpos monoclonales, métodos inmunológicos, biochips para el diagnóstico, clonación molecular y clonación animal, farmacogenes, ingeniería de proteínas, proteínas recombinantes de interés diagnóstico y terapéutico, sondas de hibridación, terapia celular y terapia génica, transferencia de genes, vectores, virus, etc. Nuestro objetivo no es describir los aspectos concretos de estos campos, sino sólo poner de manifiesto las bases moleculares que les sirven de soporte, como punto de partida para un estudio detenido de cada campo en su correspondiente área de conocimiento. Con este fin se presentan en este primer tema las características generales del genoma de eucariotas.

Experimento de Avery, MacLeod y McCarty (1944)



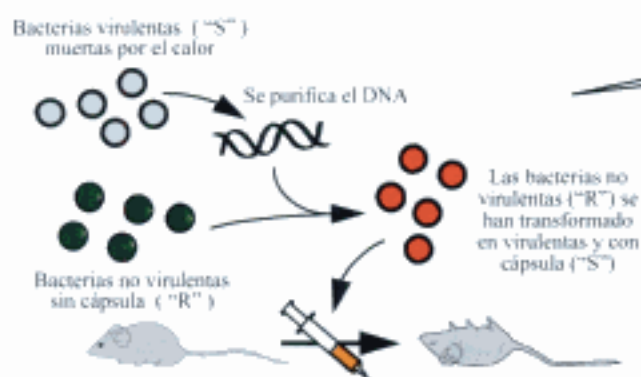
Existen cepas de neumococo (*Streptococcus pneumoniae*) rodeadas por una cápsula de polisacárido: bacterias lisas o "S" (del inglés *smooth*). Al inyectarlas producen neumonía y ocasionan la muerte a los ratones



Otras cepas de neumococo, por falta de una enzima, no producen la cápsula y son inofensivas, no patógenas: bacterias rugosas o "R" (*rough* en inglés)



Al someter cepas infecciosas "S" al calor, las bacterias mueren y dejan de ser infecciosas (Fred Griffith, 1928)

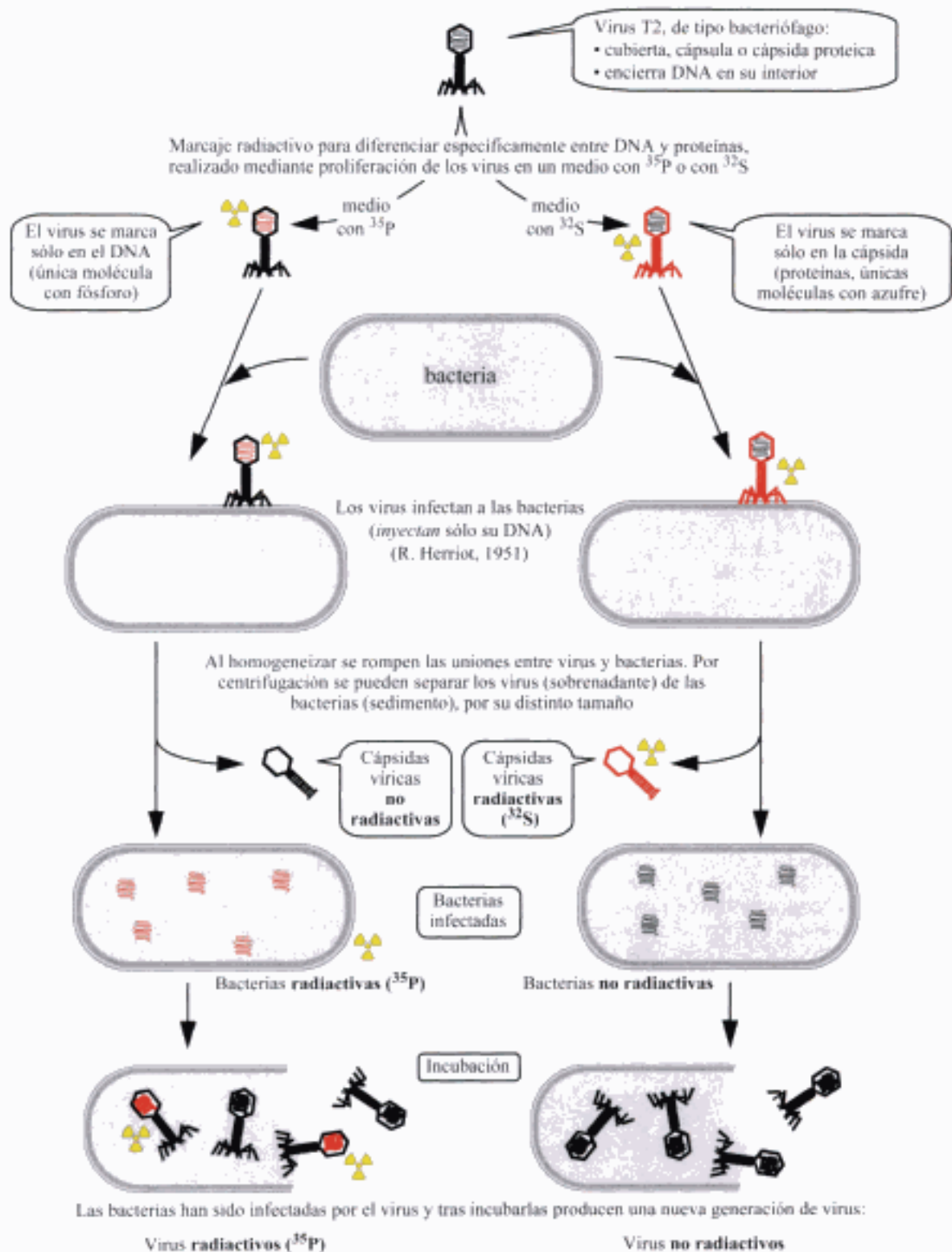


Las cepas no infecciosas "R" pueden transformarse permanentemente en infecciosas ("S") por contacto con cepas "S" o con un extracto acelular de éstas. Ese extracto se fraccionó hasta encontrar un componente con la actividad transformante, el cual poseía una composición química similar a la del DNA y propiedades físico-químicas idénticas (ópticas, ultracentrifugación, difusión, electroforesis...). Además, no se perdía la actividad al aplicar métodos de extracción de proteínas o lípidos, ni se veía afectada por el tratamiento con tripsina o quimotripsina (proteasas) o con ribonucleasa (que degrada el RNA), pero sí se perdía la actividad transformadora tras el tratamiento con desoxirribonucleasa.

En conclusión: el principio transformador era DNA.

Conclusión: El DNA es el responsable de la síntesis de la enzima necesaria para la formación de la cápsula y, por tanto, es el portador de la información genética

Experimento de Hershey y Chase (1952)



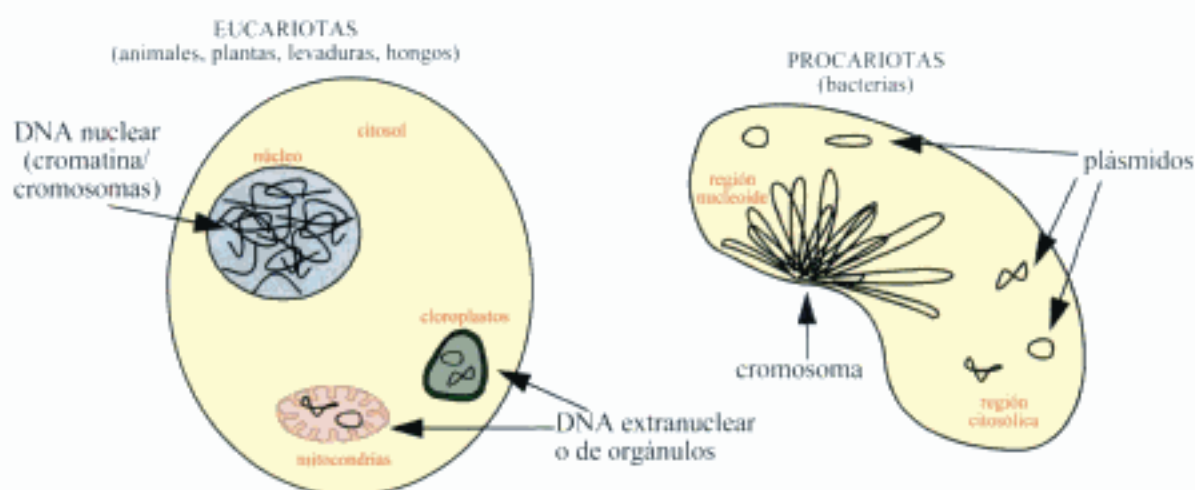
Conclusión: El DNA introducido en la bacteria contiene la información necesaria para generar nuevos virus; es decir, el DNA es el portador de la información genética.

1.3 CARACTERÍSTICAS GENERALES DEL GENOMA

1.3.1 Genoma de procariotas y eucariotas

Conceptualmente, el genoma (de la raíz griega *gen*, origen) puede ser considerado como el "patrimonio genético" de un organismo, es decir, el conjunto total de material genético hereditario presente en una célula, tanto nuclear como extranuclear, codificante o no. La mayoría de organismos poseen un genoma constituido por DNA; excepcionalmente, algunos virus tienen un genoma de RNA. La parte codificante del genoma, aunque en eucariotas es minoritaria, específica en forma de genes la información necesaria para dar lugar a productos génicos, los RNAs y las proteínas. Éstos son responsables de toda la actividad celular, de sus características morfológicas y de muchas de las claves del comportamiento de un organismo, desde la embriogénesis a su crecimiento, maduración física e intelectual, estado adulto, reproducción, y todas sus funciones metabólicas y actividades fisiológicas.

Con respecto al estudio del genoma, deben resaltarse primeramente las características diferenciales de procariotas y eucariotas:



Nota: Aunque, por claridad, en la figura no se haya reflejado, el tamaño de las células eucariotas es, en general, bastante mayor comparativamente con las procariotas. Además, la cantidad de DNA que forma los cromosomas eucarióticos es muy superior a la correspondiente al cromosoma procariótico. Igualmente, los cromosomas de orgánulos son mayores que los plásmidos.

Recuérdese que los organismos **procariotas** (o **procariontes**) son células (eubacterias y arqueobacterias) de tamaño pequeño (0,2 a 5 μm diámetro), con membrana plasmática y generalmente con una pared celular rígida, y diversas estructuras superficiales (pili, flagelos, etc.), pero carecen de membrana nuclear, de orgánulos y de cualquier otro tipo de compartimentación interna. El **genoma de procariotas** es, comparativamente, de pequeño tamaño y relativamente sencillo. Se reduce a un único **cromosoma** formado por una molécula circular de DNA, localizado en suspensión en el citosol, en la llamada *zona nuclear* o *nucleóide*, sin membrana que la delimite. Es frecuente encontrar, además, una pequeña cantidad de material genético extra, no esencial, en forma de pequeñas moléculas circulares de DNA, llamadas **plásmidos**. Los plásmidos poseen la propiedad de replicación autónoma, es decir, sin coordinación alguna con la división celular. Juegan un papel importante, al codificar proteínas que confieren resistencia a los antibióticos u otros materiales tóxicos, y ofrecen una importante aplicación como herramientas en la tecnología del DNA recombinante, concretamente como vectores para la incorporación de DNA a células.

Las células de **eucariotas** (o **eucariontes**), tanto multicelulares (animales y plantas) como unicelulares y multicelulares simples (protozoos, hongos y algas), son más variadas y complejas. Son de mayor tamaño (10-50 μm), poseen membrana plasmática, pueden poseer *pared celular* (plantas), tienen citoesqueleto interno, y muestran una gran compartimentación citoplasmática, con diferentes tipos de orgánulos subcelulares (*mitocondrias* en animales y plantas, *cloroplastos* en plantas, *lisosomas* en animales, etc.) y otras estructuras (*retículo endoplásmico*, *complejo de Golgi*, etc.). El **núcleo** tiene como característica esencial la presencia de una *doble membrana nuclear*. En el genoma de eucariotas debe distinguirse entre el genoma nuclear y el de orgánulos. El **genoma nuclear** aparece bajo dos formas estructuralmente diferentes durante el ciclo celular: la **cromatina** (en la interfase, entre divisiones), como material filamentosamente disperso por la mayor parte del núcleo, y los **cromosomas** (durante la división), fácilmente observables como entidades morfológicas o corpúsculos independientes. Ambas formas están constituidas por las mismas moléculas lineales muy largas de DNA de doble hebra, estrechamente asociadas a proteínas. El **genoma de orgánulos** (cloroplastos y mitocondrias) está organizado en cromosomas circulares, al igual que el de procariotas.

Algunas características diferenciales del genoma humano (célula somática, diploide, $2n$)

	NUCLEAR	MITOCONDRIAL
Tamaño	6.600 Mb = 6.600.000 kb = $6,6 \cdot 10^9$ pb	16,5 kb
Tipo de DNA	Lineal, de doble hebra	Circular, de doble hebra
Número de cromosomas	23 distintos, 2 copias: 46 en total	Uno, con decenas de copias en cada mitocondria; de 200 a 10^5 copias por célula
Presencia de nucleosomas	Si	No
Proteínas asociadas	Varias histonas y no-histonas	Libre de proteínas
DNA codificante	3-5%	~ 93%
DNA repetitivo	Abundante (~ 30%)	Prácticamente nada
Número de genes	65.000 - 100.000	37
Densidad de genes	~ 1 gen cada 33-50 kb	1 gen cada 0,45 kb
Presencia de intrones en los genes	En la mayoría	Ausentes

1.3.2 Magnitud del genoma nuclear de eucariotas

El material genético contenido en el núcleo supone generalmente más del 90% del total de DNA celular. Por tanto, el primer aspecto a destacar es su **magnitud** con respecto al genoma mitocondrial y al de procariotas. A diferencia de éstos, el genoma nuclear está repartido en varios cromosomas, en número variable según la especie, generalmente muy grandes y con el DNA muy condensado al estar estrechamente asociado a histonas y otras proteínas.

La magnitud del genoma viene determinada por la cantidad total de DNA en el conjunto de cromosomas de una célula. Ésta se puede expresar de varias formas:

- * En número de cromosomas, designado por n en organismos haploides y $2n$ en diploides, donde hay 2 copias de cada cromosoma.
- * En masa de DNA (**pg** = picogramos).
- * Como **valor C**: consiste en expresar la cantidad de DNA total en el genoma de una célula con respecto a la presente en una célula haploide de la misma especie. Como se verá, en células diploides el valor puede ser de $2C$ o $4C$, dependiendo del estadio del ciclo celular.
- * Lo más habitual es expresarlo en longitud: **pb** = pares de bases en hebra doble, o **nt** = nucleótidos en hebra sencilla. Para moléculas largas se emplean kilobases (**kb** o **kpb** = 10^3 pb) o megabases (**Mb** o **Mpb** = 10^6 pb).

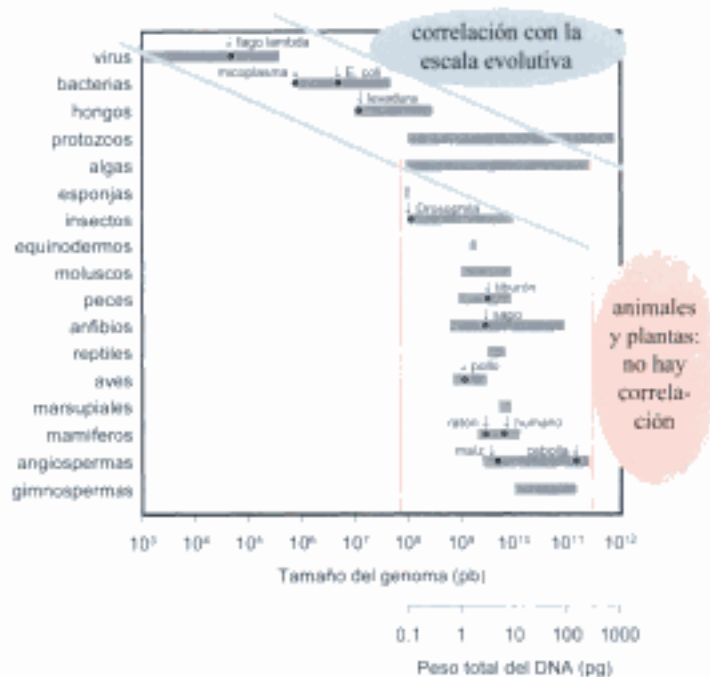
Comparaciones que resaltan la magnitud del DNA nuclear en eucariotas:

- En promedio, la cantidad total de DNA en una célula diploide eucariota es entre 8 y 200 veces superior que la de una célula procariota, aunque la diferencia puede llegar hasta 100.000 veces, por ejemplo considerando los casos extremos de algunas células de plantas y anfibios respecto al procariota de menor tamaño, el micoplasma.
- En forma lineal, la longitud total del DNA de una célula humana somática (diploide, 46 cromosomas) sería de unos **2 metros**.
- Al referir esta cifra al número de células del cuerpo humano (2 billones = $2 \cdot 10^{12}$ células), la longitud total resultante, **4 billones de metros** ($4 \cdot 10^{12}$ m), equivale a 100.000 vueltas a la Tierra (circunferencia de la Tierra = 40.000 km), 7.000 viajes de ida y vuelta a la Luna (distancia Tierra-Luna = 300.000 km) y 13 viajes de ida y vuelta al Sol (distancia Tierra-Sol = 150 millones de km).
- Para escribir la secuencia completa del genoma total de una célula humana se necesitarían 1.320 volúmenes de una obra (considerando 1.000 páginas por volumen y 5.000 letras por página), que formarían una "pila" de 260 m de altura.
- Las cifras se hacen casi incomprensibles si se quiere considerar la longitud total del DNA de toda la humanidad (actualmente $6 \cdot 10^9$ individuos).

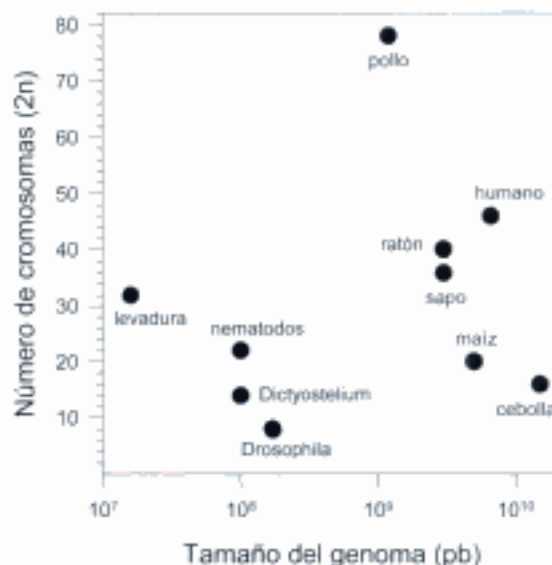
Procede, por otra parte, analizar la magnitud del genoma de forma comparativa tanto entre individuos de una misma especie como entre especies diferentes:

- Todas las células de los individuos **de una misma especie** poseen la misma cantidad total de DNA y número de cromosomas. De esta generalización deben exceptuarse las células germinales, dedicadas a la reproducción del individuo, que poseen la mitad. En todos los casos, estas cifras se refieren a células en un mismo estado de división celular (normalmente la interfase), pues, como se verá posteriormente (tema 9), la cantidad total de DNA, aunque no el número de cromosomas, varía en una célula a lo largo del ciclo celular.

b) El contenido total de DNA, expresado en pg o en pb, y el número de cromosomas varían ampliamente **entre especies diferentes**. La magnitud del genoma está relacionada con el grado de información requerido por cada especie; es decir, es mayor cuanto más complejo es el organismo, cuanto más se asciende en la escala evolutiva. Sin embargo, esta correlación sólo es válida al comparar especies muy alejadas evolutivamente. Así, por ejemplo, la célula humana contiene 700 veces más DNA que la de *E. coli*, pero 30 veces menos que algunas células de anfibios y plantas.



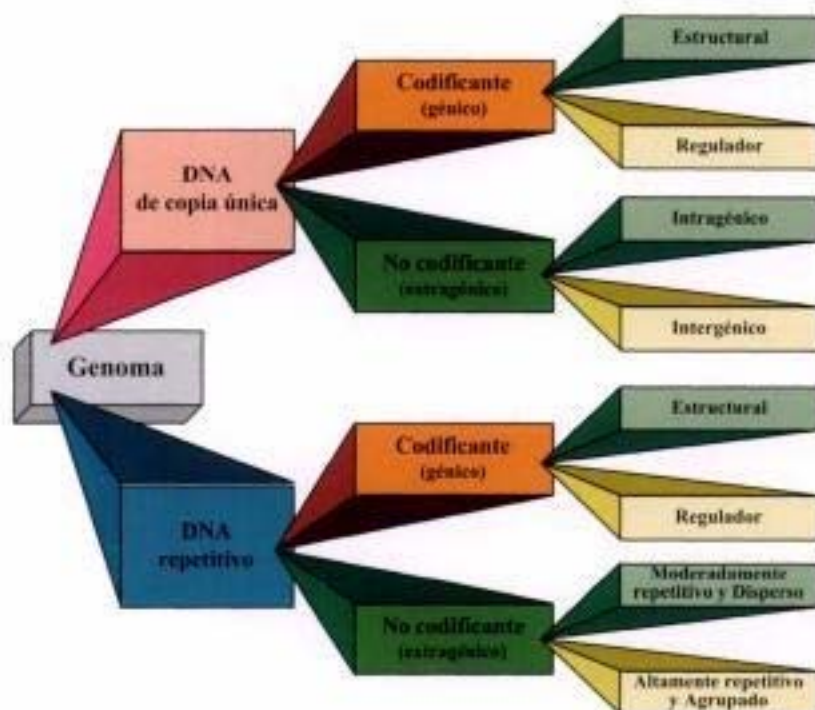
Por otro lado, no existe correlación alguna entre el tamaño del genoma de distintas especies y el número de cromosomas que lo componen. Por ejemplo, mientras que el genoma humano diploide ($6.6 \cdot 10^9$ pb) está repartido en 46 cromosomas (23 pares, $n=23$), las células de ratón, con un genoma menor ($3 \cdot 10^9$ pb), poseen menor número de cromosomas (40); por el contrario, en las células de cebolla un genoma 2 veces mayor que el humano ($1.5 \cdot 10^{10}$ pb) está contenido en sólo 16 cromosomas. De forma similar, la mosca de la fruta, con un genoma ($1.7 \cdot 10^8$ pb) mucho mayor que el de la levadura ($1.6 \cdot 10^7$ pb), posee un número muy inferior de cromosomas (8 frente a 32).



Esta ausencia de correlación entre cantidad de material genético y complejidad del organismo sugiere que el tamaño de los genomas de organismos superiores es muy superior al necesario y, por tanto, que una gran parte es *no codificante*, es decir, no está organizado en genes, nunca traduce su información a un producto génico (RNA o proteína). Esta hipótesis ha sido confirmada al estudiar con detalle el genoma. Por ejemplo, se ha encontrado que los genes tienen una longitud media de alrededor de 1.000 pb. Asumiendo esta cifra, el genoma humano haploide podría contener, en teoría, unos 3,3 millones de genes (pues cada aminoácido está codificado por 3 nucleótidos) y, sin

embargo, se estima hoy en día que sólo contiene unos 100.000, de modo que la proporción codificante sólo alcanza alrededor del 3% del DNA total.

Aunque la función de los genes está bien establecida para una pequeña fracción del total presente en el genoma humano, la inexistencia de una correlación entre cantidad total de DNA en el genoma y grado de complejidad del organismo superior puso de manifiesto la necesidad de estudiar con mayor precisión la organización estructural, a nivel molecular, tanto del genoma (tema 10) como de los genes (tema 19). Los hechos más característicos en relación con la complejidad del genoma nuclear son la presencia de grandes regiones que no codifican producto génico alguno (*DNA no codificante*) y la existencia de secuencias de DNA que se repiten un número más o menos elevado de veces (*DNA repetitivo*). Como se estudiará con detenimiento posteriormente, hay que unir a ello la existencia de una enorme diversidad de secuencias (polimorfismo, tema 26) en distintas regiones del DNA, tanto sencillas como repetitivas, codificantes o no.



1.3.3 Magnitud y características del genoma mitocondrial

El genoma de orgánulos (mitocondrias en animales, mitocondrias y cloroplastos en plantas) es, posiblemente, un vestigio del cromosoma de bacterias arcaicas que accedieron al citoplasma de eucariotas primitivos, para dar lugar tras la evolución a dichos orgánulos. Se cree que la mitocondria (orgánulo que aporta el 90% de la energía que necesitan las células) surgió al acumularse el oxígeno en la atmósfera terrestre. Posiblemente, la mitocondria y el núcleo de la célula eucariota se crearon a la vez, al incorporarse por endocitosis células procariotas aeróbicas al interior de la célula eucariótica anaeróbica y fusionarse ambas. Con el tiempo, la mayoría de los genes procarióticos (genes protomitochondriales) se integraron en el genoma nuclear, con lo que el eucariota primitivo anaeróbico ya podía vivir en una atmósfera aeróbica, rica en oxígeno; sólo una pequeña fracción del genoma procariótico primigenio permaneció en la mitocondria.

Las células animales poseen un número muy variable de mitocondrias, dependiendo del tejido. Cada mitocondria posee varias copias de un único cromosoma, situadas en la matriz mitocondrial y ancladas a la membrana interna. En consecuencia, el número de copias de DNA mitocondrial (mtDNA) en una célula oscila entre 200 y 100.000. El cromosoma mitocondrial es bicatenario y circular, como el de procariotas, aunque de tamaño muy inferior: 16.569 pb en humanos, con una longitud de 5 μm y una masa molecular de 10 MDa; esto supone unas 3.000 veces menos que la longitud del cromosoma nuclear más pequeño. Teniendo en cuenta el número de copias totales (2 del genoma nuclear y múltiples, y muy variables, del mitocondrial), se puede calcular que el DNA mitocondrial total supone, dependiendo del tejido, entre un 0,05 y un 20% del DNA total de la célula. Finalmente, el cromosoma de cloroplastos es igualmente bicatenario y circular, pero de mayor tamaño que el mitocondrial.

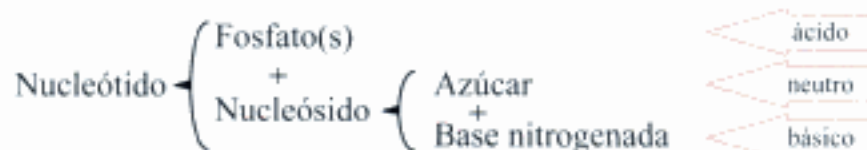
A diferencia del DNA nuclear, de gran complejidad, con regiones codificantes y no codificantes, el mtDNA es en su casi totalidad codificante y no repetitivo. En concreto, el cromosoma mitocondrial humano contiene un total de 37 genes: 2 para rRNA, 22 para tRNA y 13 para proteínas que forman parte de los complejos enzimáticos respiratorios I, III, IV y V. Éstas suponen sólo el 5% de las proteínas mitocondriales, mientras que el resto están codificadas por DNA nuclear.

Tema 2

Componentes fundamentales de los ácidos nucleicos

2.1	Componente ácido: Fosfatos.....	9
2.1.1	Monofosfatos.....	9
2.1.2	Difosfatos.....	10
2.1.3	Trifosfatos.....	10
2.1.4	Tipos de enlaces fosfato.....	11
2.2	Componente neutro: Azúcares.....	11
2.3	Componente básico: Bases nitrogenadas.....	12
2.3.1	Bases nitrogenadas más frecuentes: 2 tipos.....	12
2.3.2	Otras bases de interés biológico y clínico.....	13
2.3.3	Propiedades fisicoquímicas de las bases purínicas y pirimidínicas.....	14
2.3.3.1	Existencia de dipolos.....	14
2.3.3.2	Hidrofobicidad.....	14
2.3.3.3	Disposición coplanar de los enlaces de cada anillo (C-N y C-C).....	14
2.3.3.4	Tautomería o isomería dinámica.....	14
2.3.3.5	Carácter básico.....	16
2.3.3.6	Absorción de la luz en el ultravioleta.....	17

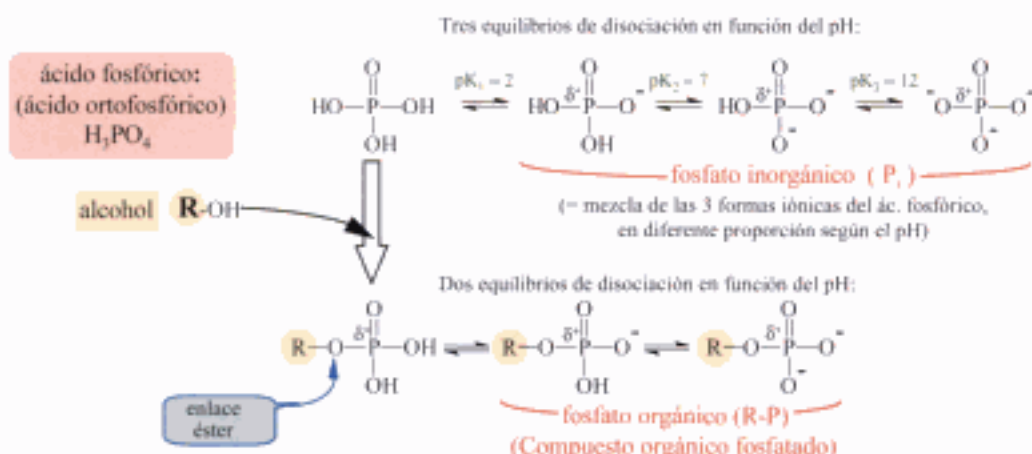
Los ácidos nucleicos (DNA y RNA) están constituidos por la unión de numerosos nucleótidos, de 4 tipos. Como introducción a su estudio, se consideran algunos aspectos de interés para cada uno de los componentes estructurales de dichas unidades, los nucleótidos:



2.1 COMPONENTE ÁCIDO: FOSFATOS

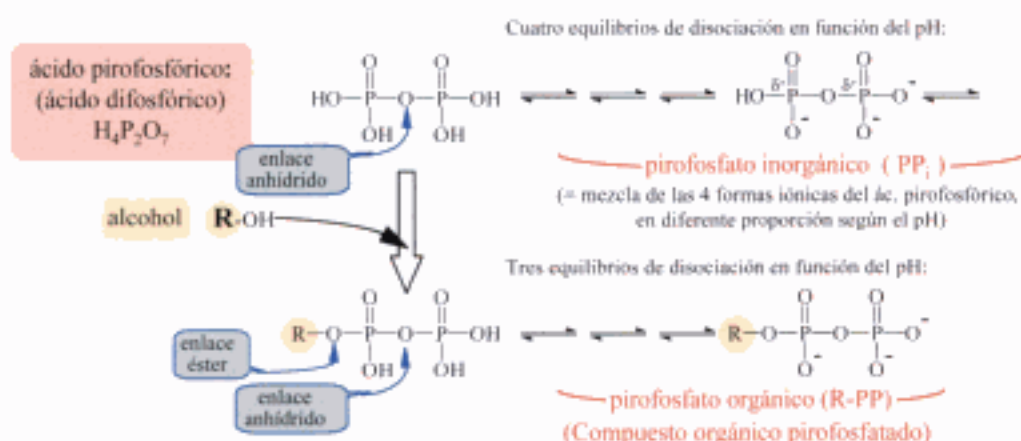
2.1.1 Monofosfatos

El **fosfato inorgánico** (P_i), mezcla de las especies químicas resultantes de la disociación de la molécula de ácido fosfórico, aparece como anión a pH ácido, como dianión a pH fisiológico y como trianión a pH alcalino. En la célula, el fosfato se encuentra en forma libre (P_i) o formando parte de numerosos compuestos orgánicos (R-P), por ejemplo, los nucleótidos.



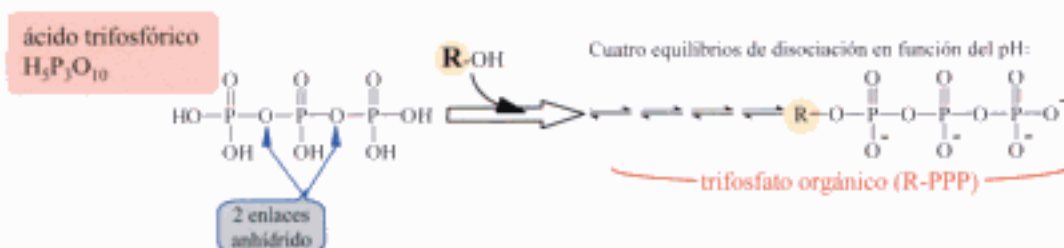
2.1.2 Difosfatos

El **difosfato o pirofosfato (PP_i)** se presenta a pH fisiológico como trianión (una de las 4 formas iónicas del ácido pirofosfórico). Es estable en disolución a pH neutro o alcalino; sin embargo, dentro de la célula es inestable por acción de las pirofosfatasas, formando 2 P_i . Aparece combinado fundamentalmente en nucleósidos-difosfato (por ejemplo, ADP).



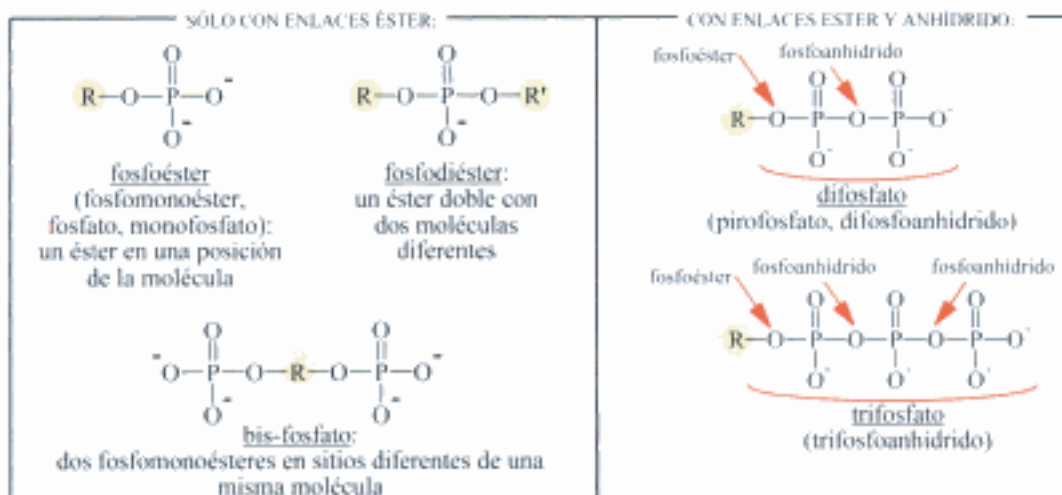
2.1.3 Trifosfatos

El **trifosfato (PPP_i)** se presenta a pH fisiológico como tetraanión. No aparece en forma libre, sino formando parte, esencialmente, de nucleósidos-trifosfato (por ejemplo, ATP).



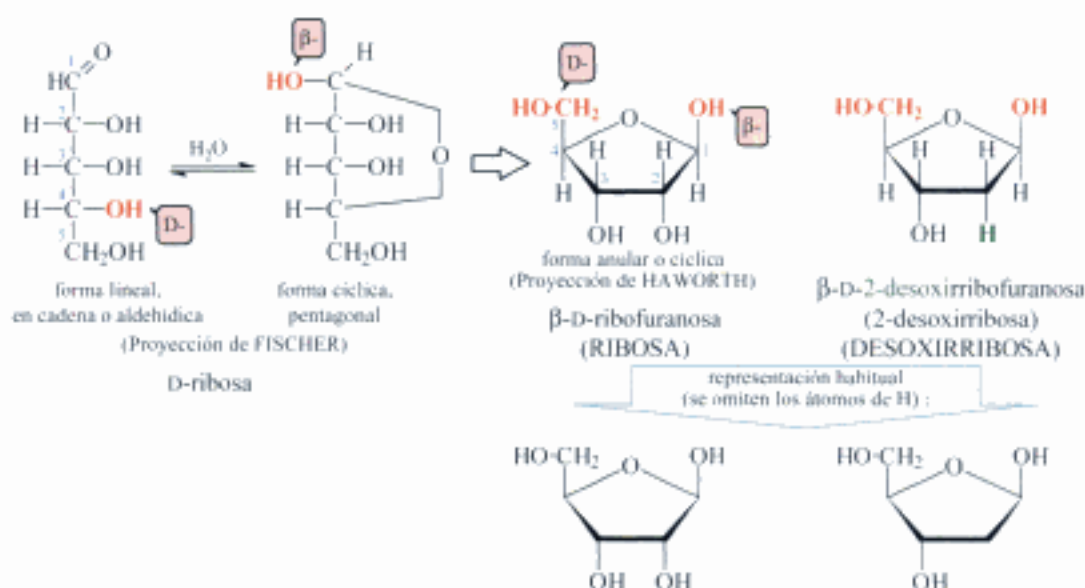
2.1.4 Tipos de enlaces fosfato

Pueden distinguirse cinco **tipos de enlaces fosfato** que van a formar parte de los nucleótidos. Obsérvese la estructura y nombre de todos ellos, así como la diferencia entre enlace éster y anhídrido en los nucleótidos difosfatados y trifosfatados.



2.2 COMPONENTE NEUTRO: AZÚCARES

El azúcar (inicialmente llamado *osa*) siempre es una pentosa (5 átomos de carbono), bien la D-ribosa o la D-desoxirribosa. La **ribosa** (nombre derivado de **osa** del Rockefeller Institute of Biochemistry, 1908) interviene en nucleósidos y nucleótidos bajo su forma más estable, β -D-ribofuranosa.

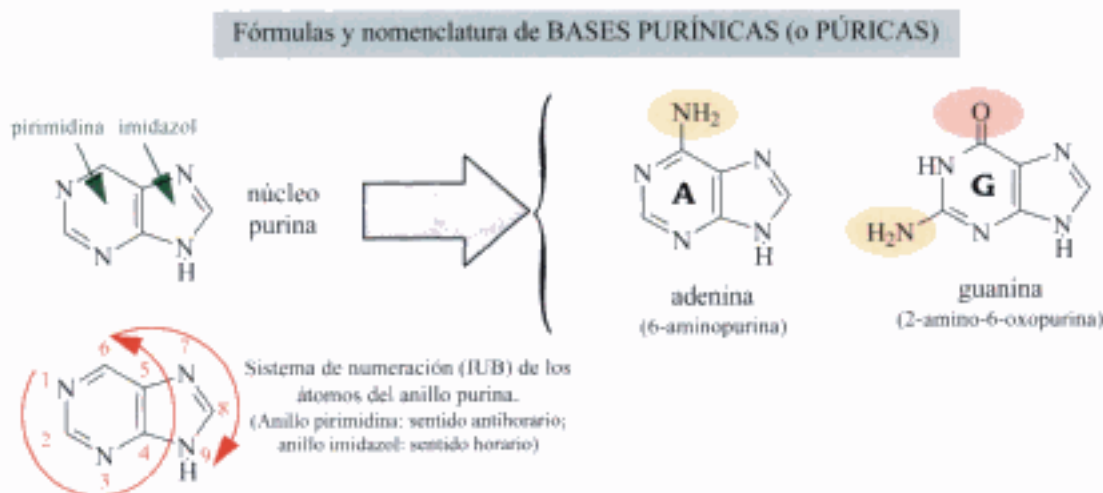
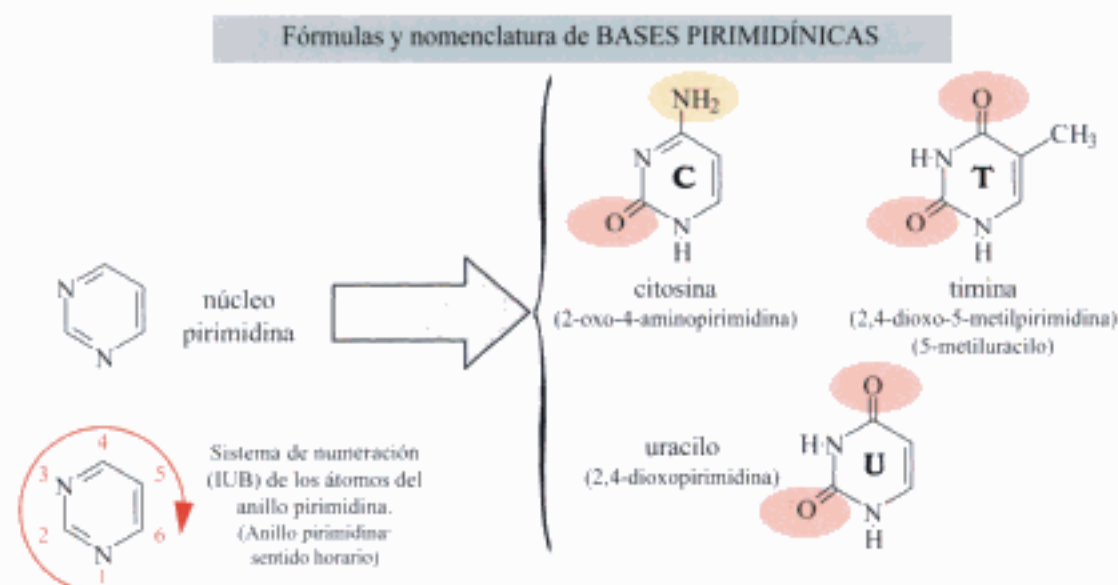


La fórmula representada como ejemplo puede extenderse a otras muchas moléculas de interés bioquímico: D-glucosa, D-galactosa, 3-desoxi-D-ribosa, etc.

2.3 COMPONENTE BÁSICO: BASES NITROGENADAS

Son moléculas heterocíclicas, con más de un átomo de nitrógeno, formadas por un anillo único (pirimidina) o por dos anillos condensados (purina). Obsérvense en los ejemplos de bases que intervienen en los ácidos nucleicos la numeración de los átomos del heterociclo y la posición de los sustituyentes externos.

2.3.1 Bases nitrogenadas más frecuentes: 2 tipos

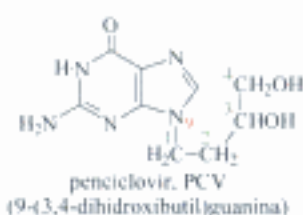
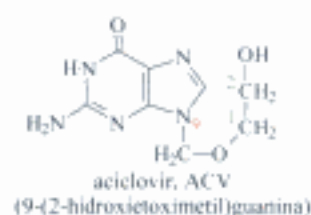
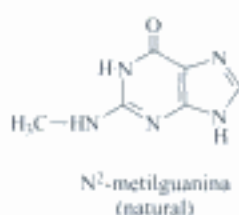
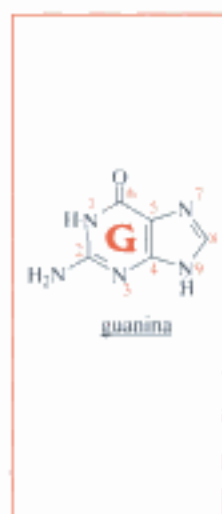
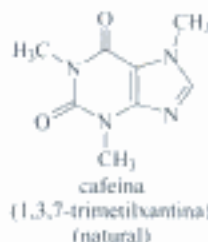
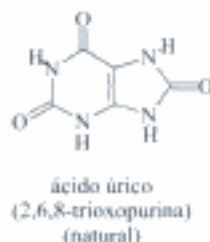
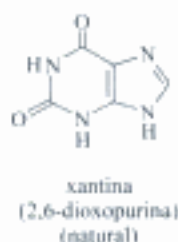
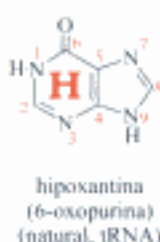
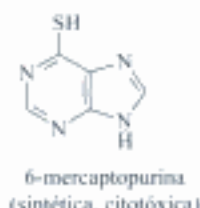
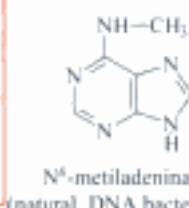
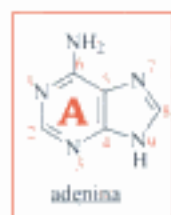


A modo de resumen, la posición de los sustituyentes presentes en cada una de las bases nitrogenadas es:

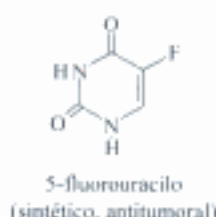
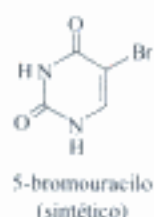
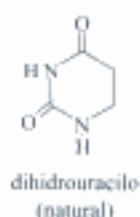
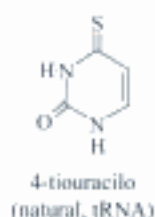
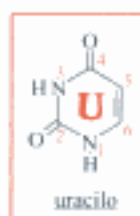
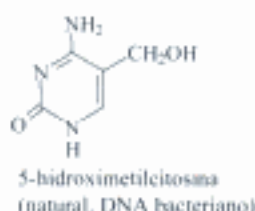
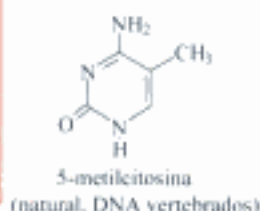
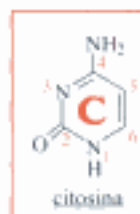
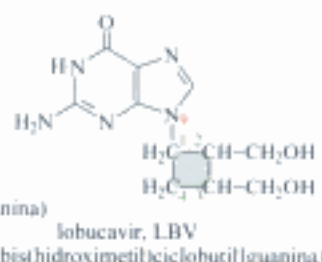
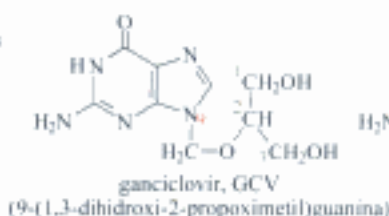
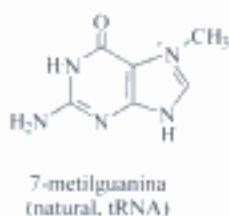
	-NH ₂	-O	-CH ₃
A	6		
C	4	2	
U		2 y 4	
G	2	6	
T		2 y 4	5

2.3.2 Otras bases de interés biológico y clínico

Existen otras bases purínicas y pirimidínicas, en general poco frecuentes, que surgen por modificación de las bases principales A, G, C y U o de alguno de sus análogos (hipoxantina, H). Son de origen natural o sintético, intervienen en procesos como la metilación del DNA en la célula y muestran distintas propiedades (antibióticas, antivíricas, antitumorales, etc.).



Análogos sintéticos usados en terapia antiviral



2.3.3 Propiedades fisicoquímicas de las bases purínicas y pirimidínicas

2.3.3.1 Existencia de dipolos

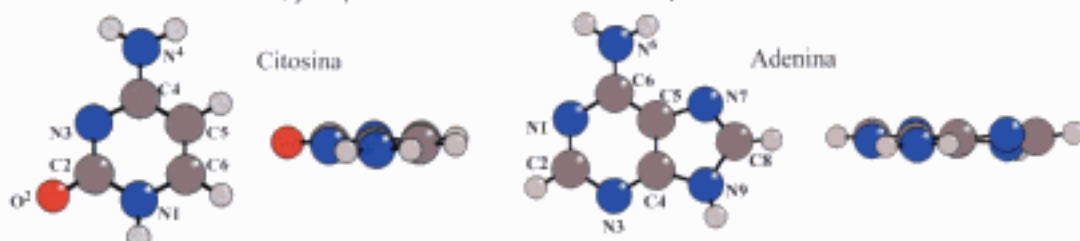
Todas las bases poseen átomos electronegativos de O (excepto adenina), en posición exocíclica o extranuclear, y de N, tanto exocíclicos como en el anillo, nucleares. Como consecuencia, son abundantes los enlaces polares, lo que les permite interactuar entre sí mediante puentes de hidrógeno, que poseen una importancia capital en la estructura secundaria de los ácidos nucleicos, en especial del DNA.

2.3.3.2 Hidrofobicidad

La naturaleza aromática de los anillos hace que las bases tengan un marcado carácter apolar, sean hidrofóbicas y poco solubles en agua al pH celular, cercano a la neutralidad. A pH ácido o alcalino adquieren carga y se hacen más solubles en agua. Como ya se estudiará, las **interacciones hidrofóbicas de apilamiento**, que tienen lugar entre bases dispuestas paralelamente (a modo de monedas apiladas), son esenciales para estabilizar la estructura tridimensional de los ácidos nucleicos.

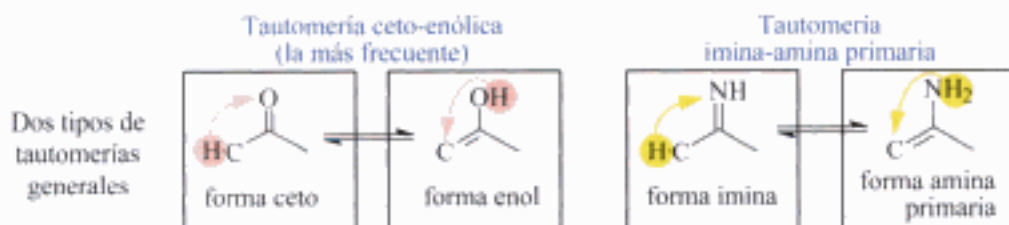
2.3.3.3 Disposición coplanar de los enlaces de cada anillo (C-N y C-C)

Se debe al carácter aromático, y se puede observar en modelos espaciales:

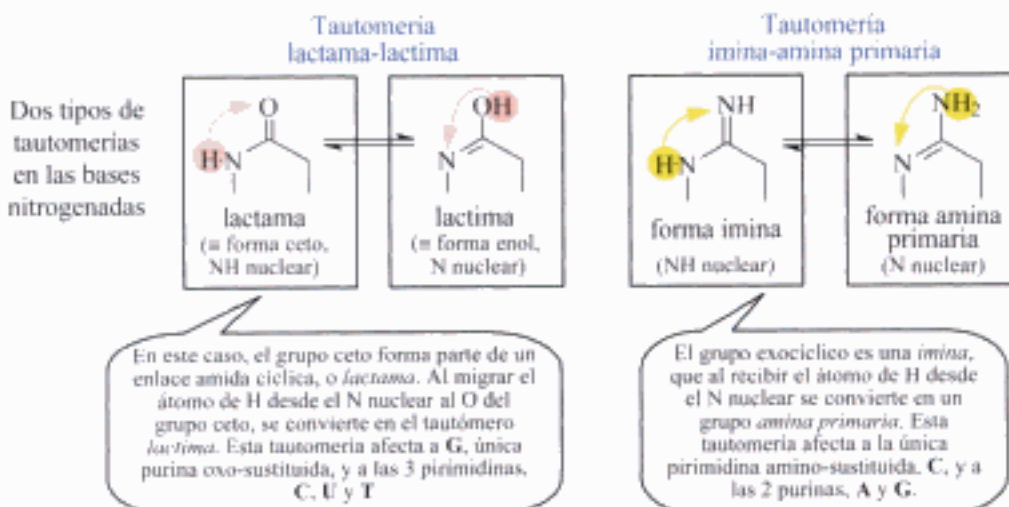


2.3.3.4 Tautomería o isomería dinámica

Es una isomería de grupos funcionales. Consiste, con carácter general, en la migración del átomo de hidrógeno unido a un átomo de C, O o N hacia otro átomo adyacente. El tipo más frecuente es la tautomería ceto-enólica.



Son muchos los compuestos orgánicos que existen como mezclas de 2 o más de estas **formas tautoméricas**, interconvertibles y en equilibrio. Entre los compuestos heterocíclicos son de especial importancia las tautomerías en las bases purínicas y pirimidínicas, en las cuales la migración del H tiene lugar desde un átomo de N nuclear hacia otro átomo extranuclear, y viceversa:

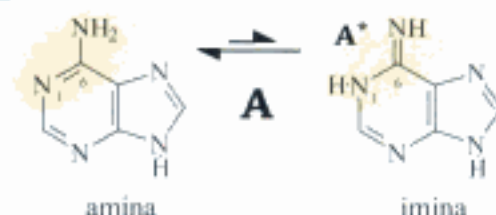


En consecuencia, las tautomerías de las bases nitrogenadas son las siguientes:

Una base (A) presenta una sola tautomería:

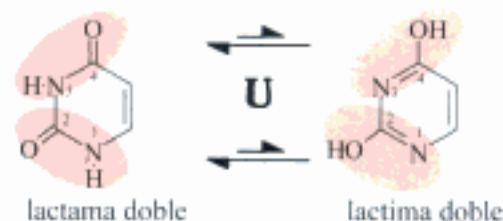
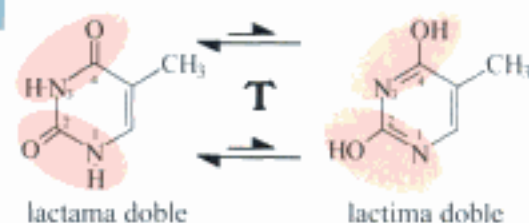
La tautomería de A es imina-amina primaria.

Los tautómeros imina se nombran poniendo un asterisco sobre la letra de la base.



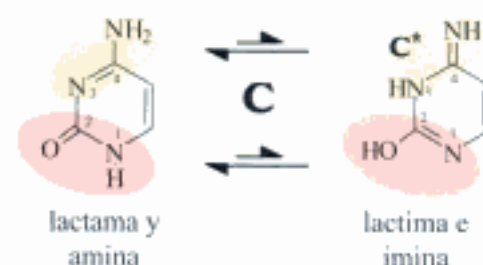
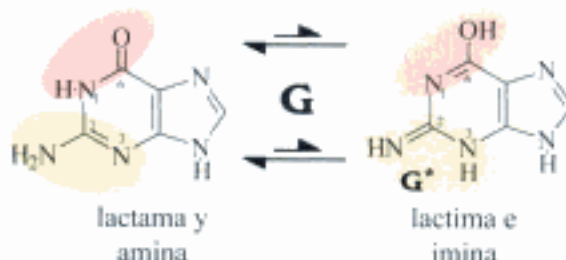
Cuatro bases (T, U, G, C) presentan cada una dos tautomerías independientes:

T y U muestran tautomería doble lactama-lactima



G y C muestran tautomerías imina-amina primaria y lactama-lactima.

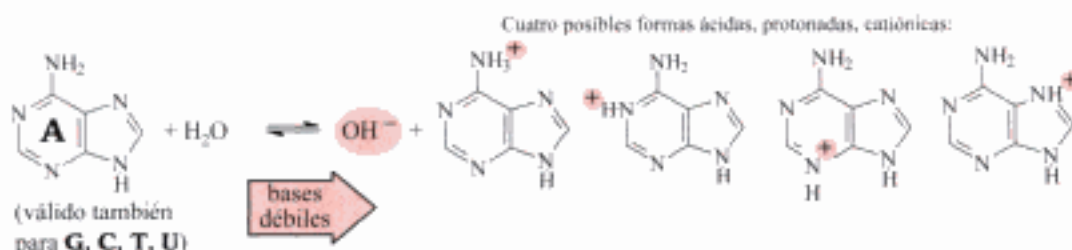
Los tautómeros imina se nombran poniendo un asterisco sobre la letra de la base.



Desde el punto de vista químico, en las bases libres las formas tautoméricas lactima (enol) e imina son menos estables, y por ello son unas 10.000 veces menos abundantes que sus correspondientes tautómeros lactama (ceto) y amina, con lo que el equilibrio se encuentra desplazado hacia la izquierda.

2.3.3.5 Carácter básico

Todas las bases nitrogenadas son bases débiles, pues sus átomos de N, tanto nucleares como extranucleares, se pueden protonar, con valores de pK_a comprendidos entre 9 y 10. Como ejemplo representativo, obsérvese la ionización de la adenina:



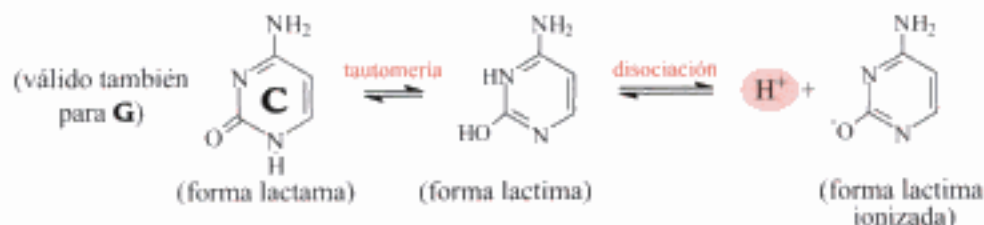
Sin embargo, este carácter básico, propio de los anillos purina y pirimidina, se ve disminuido por la presencia de algunos sustituyentes. En concreto, el grupo $-OH$ de los enoles (y lactimas) es ligeramente ácido y puede disociarse:



Como consecuencia, las bases nitrogenadas que poseen grupos lactama (ceto) pueden actuar como ácidos, a través de sus tautómeros lactima (enol), con lo cual se ve reducido el carácter básico propio de los anillos, en mayor medida cuanto mayor sea el número de grupos $-OH$. Véanse los ejemplos siguientes:

a) **A** es la única base que no tiene grupo ceto/enol y, por tanto, no muestra ningún carácter ácido y es la más básica de todas.

b) En **C** y **G**, el único grupo OH aporta un carácter ácido muy débil, por lo que siguen comportándose como bases:



c) En **T** y **U**, la presencia de dos grupos OH compensa en mayor grado el carácter básico:



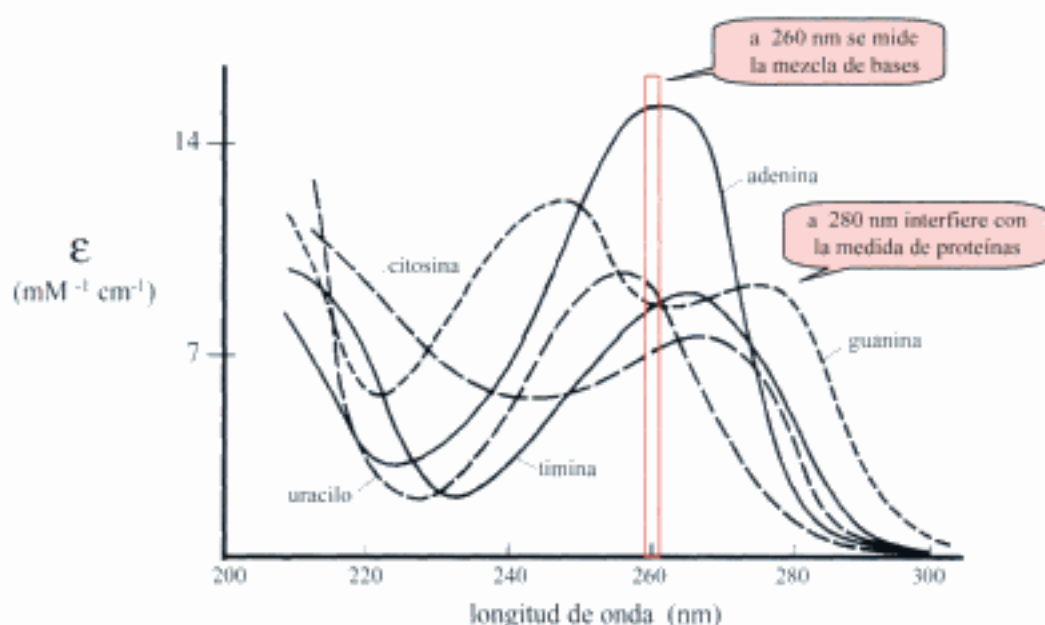
d) Como ejemplo extremo, el ácido úrico (base nitrogenada con tres grupos OH) ya muestra un marcado carácter ácido (aunque no es un ácido fuerte):



Esta disociación, como ácidos, de los grupos enol compensa en parte el desequilibrio tautomérico anteriormente considerado, aumentando la presencia de las formas lactima (enol). No ocurre un efecto similar con las formas amina, por lo que en la práctica es común que en las bases nitrogenadas sólo se describa la tautomería lactama-lactima, de mayor trascendencia biológica.

2.3.3.6 Absorción de la luz en el ultravioleta

Es una propiedad característica de las bases purínicas y pirimidínicas, debida a su carácter **aromático**. En los espectros de absorción, a pH=7, se observa una fuerte absorción de la luz en el ultravioleta (UV), con un máximo cerca de 260 nm para todas ellas.



El mantenimiento de esta propiedad en nucleósidos, nucleótidos y ácidos nucleicos, de los que las bases forman parte, permite su utilización analítica en estos compuestos; por ejemplo, para la cuantificación de ácidos nucleicos (pág. 36).

Tema 3

Nucleósidos y nucleótidos

3.1 Estructura de nucleósidos	19
3.1.1 Características generales	20
3.1.2 Tipos de nucleósidos	20
3.1.2.1 Nucleósidos que forman parte de ácidos nucleicos	20
3.1.2.2 Otros nucleósidos de interés biológico y clínico	21
3.1.2.3 Análogos de nucleósidos, inhibidores de la transcriptasa inversa	23
3.2 Estructura de nucleótidos	23
3.2.1 Características generales	23
3.2.2 Nucleótidos sencillos	24
3.2.2.1 Nomenclatura	25
3.2.2.2 Estructura de los nucleósidos-monofosfato que forman parte de los ácidos nucleicos	27
3.2.3 Propiedades fisicoquímicas de nucleósidos y nucleótidos	27
3.2.3.1 Propiedades iónicas	27
3.2.3.2 Formación de complejos	27
3.2.3.3 Absorción de la luz en el ultravioleta	28
3.2.3.4 Nucleótidos como moléculas de reserva energética	28
3.2.3.5 Hidrólisis	28
3.2.4 Nucleótidos cíclicos	29
3.2.5 Nucleótidos complejos	29

3.1 ESTRUCTURA DE NUCLEÓSIDOS

Corresponde al esquema general: **nucleósido = base nitrogenada + azúcar**

También pueden considerarse como el producto de la hidrólisis de los nucleótidos, con liberación de fosfato inorgánico. Equivale, por tanto, a un nucleótido carente de fosfatos.



3.1.1 Características generales

La unión covalente (enlace N-glicosídico) entre la base y el azúcar (ribosa o desoxirribosa) se establece entre el C-1' (carbono anomérico) del azúcar y el N-1 del anillo pirimidínico o el N-9 del anillo purínico, con pérdida de una molécula de agua.



3.1.2 Tipos de nucleósidos

Debe distinguirse entre nucleósidos constituyentes de ácidos nucleicos y otros nucleósidos de interés biológico y clínico. Su descripción siempre se hace en función de la naturaleza química de la base y del azúcar:

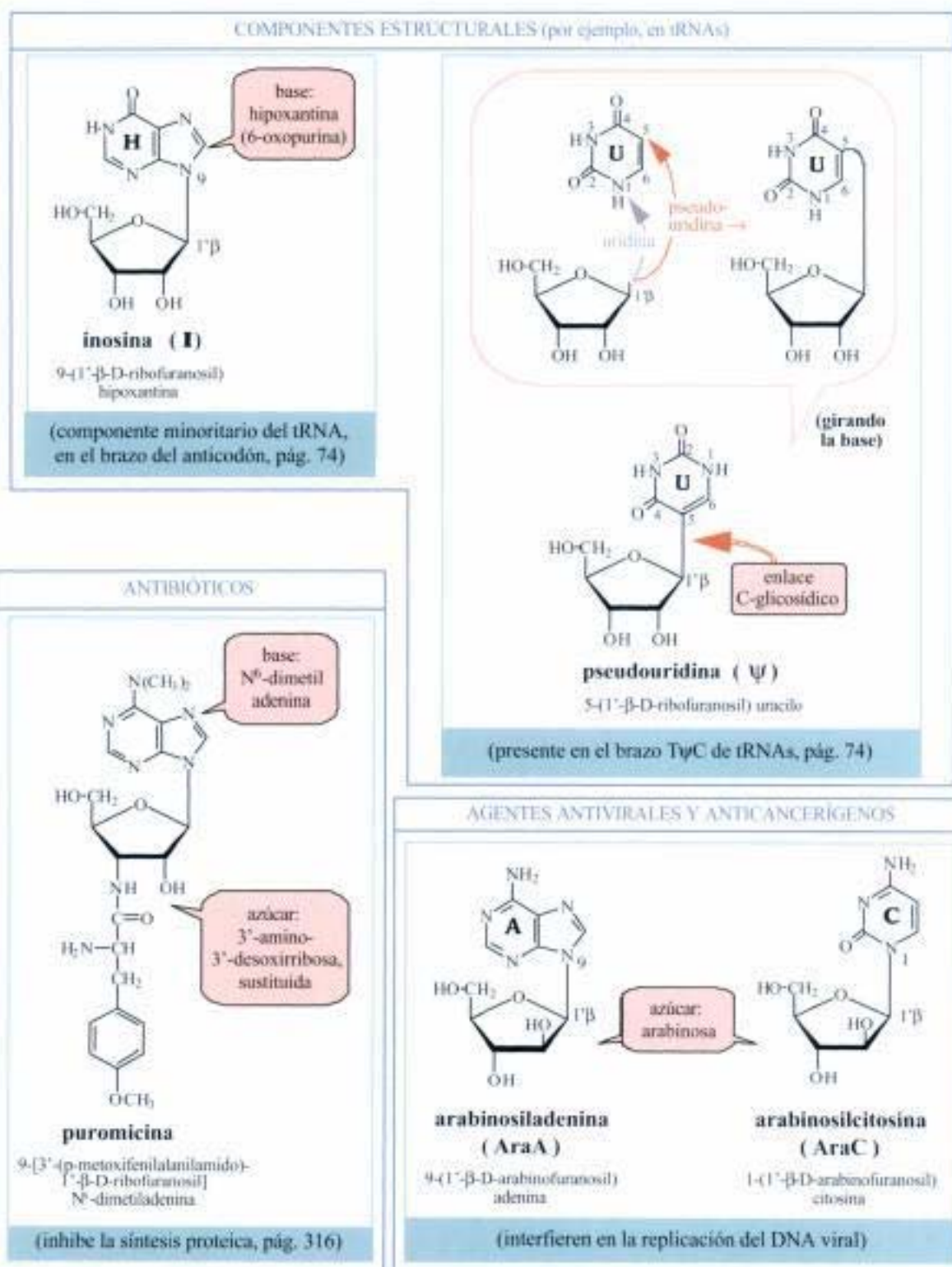
3.1.2.1 Nucleósidos que forman parte de ácidos nucleicos

Nucleósidos purínicos		Nucleósidos pirimidínicos			
Tipo	ribonucleósidos (se encuentran en el RNA) y desoxirribonucleósidos (se encuentran en el DNA)		ribonucleósido (en RNA) y desoxirribonucleósido (en DNA)	desoxirribonucleósido (en DNA)	ribonucleósido (en RNA)
Estructura					
Nombre	sufijo -osina		sufijo -idina		
	adenosina 9-(1'-β-D-ribofuranosil) adenina	guanosina 9-(1'-β-D-ribofuranosil) guanina	citidina 1-(1'-β-D-ribofuranosil) citosina	timidina (no existe en RNA, excepcionalmente en DNA)	uridina 1-(1'-β-D-ribofuranosil) uracilo
	2'-desoxiadenosina 9-(2'-desoxi-1'-β-D-ribofuranosil) adenina	2'-desoxiguanosina 9-(2'-desoxi-1'-β-D-ribofuranosil) guanina	2'-desoxicitidina 1-(2'-desoxi-1'-β-D-ribofuranosil) citosina	2'-desoxitimidina (coloquial: timidina) 1-(2'-desoxi-1'-β-D-ribofuranosil) timina	— 2'-desoxiuridina (no existe en DNA)

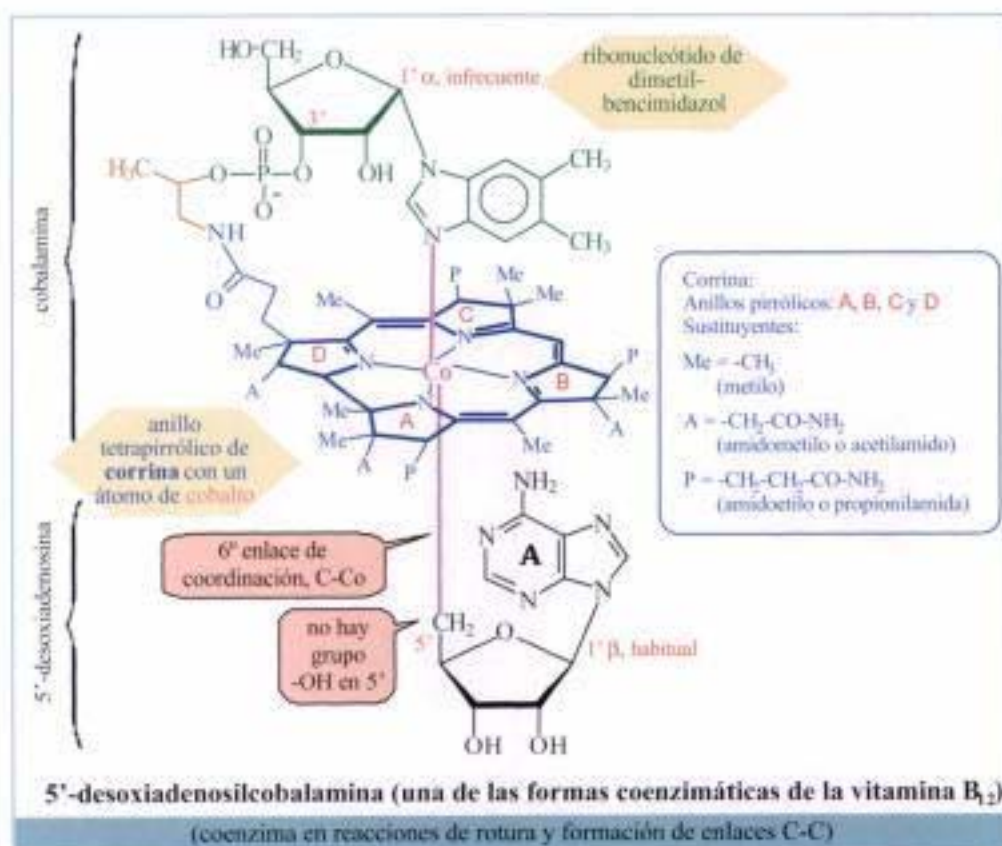
La distinción entre ribonucleósidos y desoxirribonucleósidos estriba, obviamente, en la presencia de ribosa o desoxirribosa en la molécula.

3.1.2.2 Otros nucleósidos de interés biológico y clínico

Además de los componentes de nucleótidos y ácidos nucleicos, existen en la naturaleza otros nucleósidos en forma libre o combinada. Algunos de ellos son significativos por su localización o acción bioquímica, e incluso por sus aplicaciones clínicas. Deben citarse los siguientes:



COMPONENTES DE COENZIMAS



3.1.2.3 Análogos de nucleósidos, inhibidores de la transcriptasa inversa

Se emplean para combatir los retrovirus de la inmunodeficiencia humana (VIH-1 y VIH-2), causantes del SIDA. En realidad actúan como profármacos, ya que la célula anfitriona los convierte en los correspondientes desoxinucleósidos-5'-trifosfato, que son los verdaderos fármacos, pues compiten con los desoxirribonucleótidos normales, impidiendo que la transcriptasa inversa, propia del virus, sintetice el DNA complementario al RNA vírico.

ANTIRRETROVIRALES					
Base de la que derivan	Timina		Citosina		Hipoxantina
Estructura					
Nombre completo	3'-azido-2',3'-dideoxi-timidina	2',3'-dideshidro-2',3'-dideoxi-timidina	2',3'-dideoxi-citidina	2'-desoxi-3'-tiacitidina	2',3'-dideoxi-inosina
Nombre comercial	AZT Zidovudina	ddT Estavudina	ddC Zalcitabina	3TC Lamivudina	ddI Didanosina

3.2 ESTRUCTURA DE NUCLEÓTIDOS

Corresponde al esquema general:

nucleótido = nucleósido + fosfato(s) = base nitrogenada + azúcar + fosfato(s)

También se puede considerar como:

nucleótido = azúcar-fosfato + base nitrogenada



3.2.1 Características generales

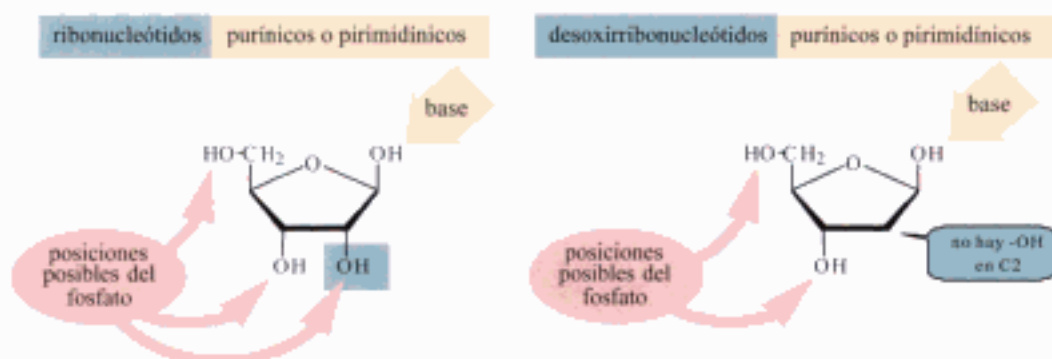
La unión, de carácter covalente (fosfomonoéster), se establece entre uno o varios grupos fosfato y sendos grupos OH del azúcar del nucleósido.

La importancia de los nucleótidos es excepcional, principalmente porque son los monómeros constituyentes de los ácidos nucleicos y, por tanto, los productos normales de su hidrólisis. Al mismo tiempo, porque aparecen en forma libre en distintas localizaciones subcelulares, formando parte de moléculas de gran interés bioquímico (especialmente coenzimas).

Para estudiar los nucleótidos se emplean distintos **criterios de clasificación**, atendiendo a:

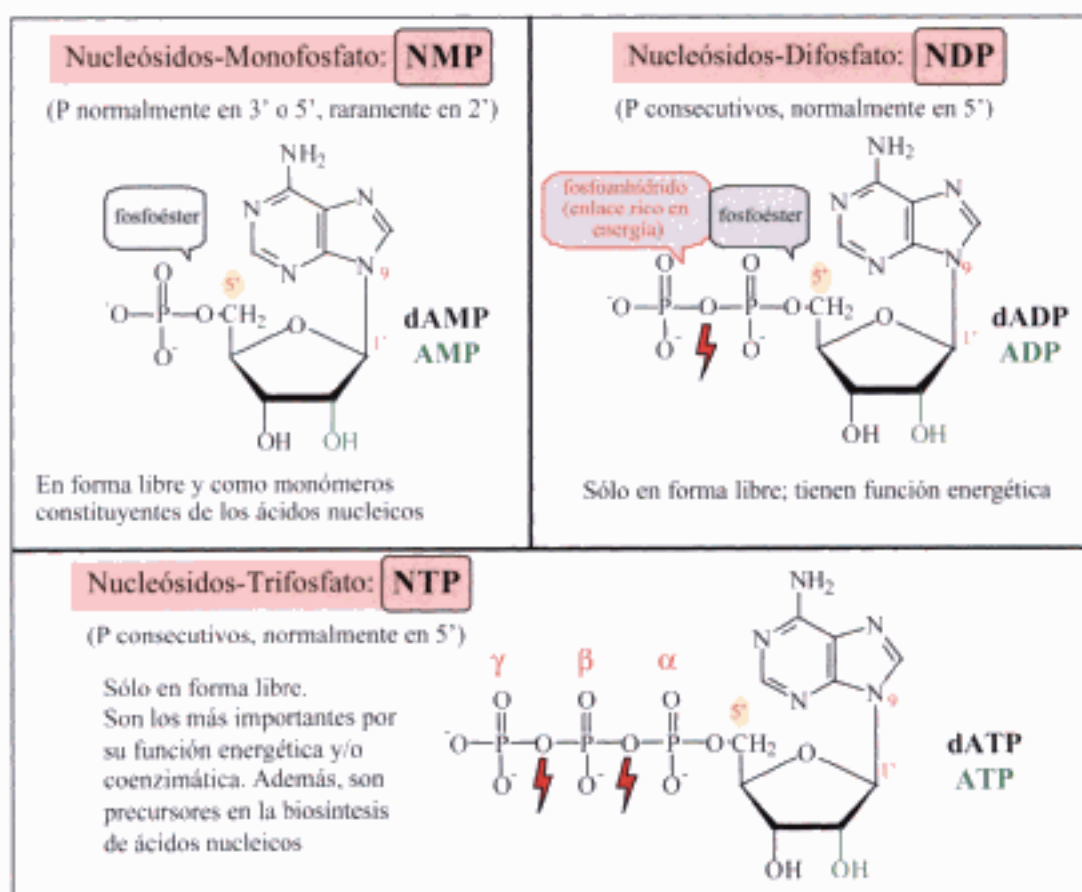
- Características estructurales, criterio utilizado normalmente:
- El número y posición en el azúcar de los grupos fosfato: nucleósidos -monofosfato, -bisfosfato y -trifosfato.
- El tipo de azúcar: ribonucleótidos y 2'-desoxirribonucleótidos.
- El tipo de base nitrogenada: nucleótidos purínicos y pirimidínicos.
- Carácter lineal o cíclico (determinado por el fosfato).

- Otras características, a las que sólo se alude ocasionalmente:
- Procedencia o estado en la célula
- Función, etc.



3.2.2 Nucleótidos sencillos

Se consideran nucleótidos sencillos los que presentan fosfato(s) en una sola posición del azúcar; esto incluye tanto los que tienen un único fosfato (*monofosfatos*) como los que tienen dos o tres fosfatos que están unidos entre sí (*difosfatos* y *trifosfatos*). Se muestran como ejemplos los correspondientes a adenina:



Los nucleótidos que forman parte de los ácidos nucleicos son siempre NMPs con un fosfato en posición 5'. Su estructura y nomenclatura se estudian de forma detallada a continuación. Los nucleótidos con 2 ó 3 fosfatos consecutivos en 5' (*difosfatos* o NDPs y *trifosfatos* o NTPs, respectivamente), dada su importante función bioquímica, son objeto de estudio en cursos de Bioquímica General. Finalmente, otros nucleótidos, con fosfatos en múltiples posiciones del azúcar, se estudian posteriormente como nucleótidos cíclicos o complejos (pág. 29).

3.2.2.1 Nomenclatura

Los nucleótidos **se nombran** según dos criterios:

a) Como **ésteres fosfato del nucleósido**. Por ej.: adenosina-monofosfato, guanosina-difosfato, uridina-trifosfato, etc. (abreviados, respectivamente, como AMP, GDP, UTP).

b) Como **ácidos**, sustituyendo la terminación del nombre de la base por el sufijo **-ílico**. Este criterio sólo se utiliza normalmente para los nucleósidos-monofosfato. Por ej.: ácido adenílico, ácido guanílico, ácido uridílico, etc. (respectivamente, para AMP, GMP y UMP). En el caso de nucleósidos-difosfato y -trifosfato, se emplearían los nombres diadenílico, triadenílico, diguanílico, etc. (para ADP, ATP y GDP). Si se quiere tener en cuenta que los restos fosfato están ionizados a pH fisiológico, se utiliza el sufijo **-ilato**: adenilato, guanilato, uridilato, etc.

Correspondencia entre los nombres de bases, nucleósidos y nucleótidos:

abreviatura común:	A	G	U	T	C
base:	adenina	guanina	uracilo	timina	citósina
nucleósido:	adenosina	guanosina	uridina	timidina	citidina
nucleótido: (nucleósido- monofosfato, NMP)	adenilato, ác. adenílico o AMP	guanilato, ác. guanílico o GMP	uridilato, ác. uridílico o UMP	timidilato, ác. timidílico o TMP	citidilato, ác. citidílico o CMP

Bajo ambos criterios, para indicar de forma más detallada la estructura, se especifica el número y la posición de los fosfatos. Por ej.:

adenosina-5'-monofosfato	ácido 5'-adenílico	5'-AMP
adenosina-3'-monofosfato	ácido 3'-adenílico	3'-AMP
adenosina-5'-difosfato	ácido 5'-diadenílico	5'-ADP
adenosina-5'-trifosfato	ácido 5'-triadenílico	5'-ATP

Cuando no se especifica el número, implícitamente se asume que el fosfato ocupa la posición 5'. En este sentido, los dos últimos ejemplos se nombran simplemente como ADP y ATP.

Estos sistemas de nomenclatura se aplican tanto a desoxirribonucleótidos como a ribonucleótidos. Lógicamente, para los primeros debe anteponerse el prefijo *desoxi-* o su forma abreviada, la letra **d**. Por ejemplo: desoxitimidina monofosfato, dTMP, desoxitimidilato, 5'-dTMP.

Una forma aún más abreviada de representar los nucleósidos monofosfato es mediante la inicial de su base, anteponiendo la letra **"p"** si el fosfato se une en 5' o posponiendo la **"p"** si la unión es en 3'. Por ej.: pA, pG, pU, Ap, Gp, Up. Este criterio es de interés para el estudio de la hidrólisis química o enzimática de los ácidos nucleicos (págs. 35 y 198).

Las tablas siguientes resumen todos los nucleótidos sencillos (NMPs, NDPs y NTPs) existentes en la célula, incluyendo sus distintas denominaciones.

NOMENCLATURA DE NUCLEÓTIDOS SENCILLOS PURÍNICOS

Bases:	Adenina		Guanina	
	Ribonucleótidos	Desoxirribo-nucleótidos	Ribonucleótidos	Desoxirribo-nucleótidos
mono-fosfato	2'-AMP 2'-adenilato adenosina-2'-fosfato 2'-adenosina monofosfato	--- (*)	2'-GMP 2'-guanilato guanosina-2'-fosfato 2'-guanosina monofosfato	--- (*)
	3'-AMP (Ap) 3'-adenilato adenosina-3'-fosfato 3'-adenosina monofosfato	3'-dAMP desoxi-etc.	3'-GMP (Gp) 3'-guanilato guanosina-3'-fosfato 3'-guanosina monofosfato	3'-dGMP desoxi-etc.
	5'-AMP (pA) 5'-adenilato adenosina-5'-fosfato 5'-adenosina monofosfato	5'-dAMP desoxi-etc.	5'-GMP (pG) 5'-guanilato guanosina-5'-fosfato 5'-guanosina monofosfato	5'-dGMP desoxi-etc.
di-fosfato	5'-ADP (ppA) adenosina-5'-difosfato 5'-adenosina difosfato	5'-dADP desoxi-etc.	5'-GDP (ppG) guanosina-5'-difosfato 5'-guanosina difosfato	5'-dGDP desoxi-etc.
tri-fosfato	5'-ATP (pppA) adenosina-5'-trifosfato 5'-adenosina trifosfato	5'-dATP desoxi-etc.	5'-GTP (pppG) guanosina-5'-trifosfato 5'-guanosina trifosfato	5'-dGTP desoxi-etc.

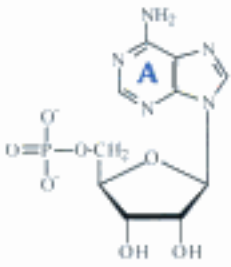
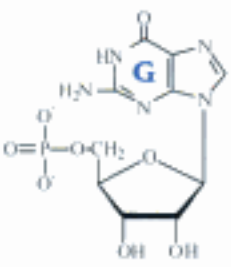
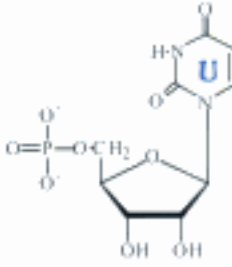
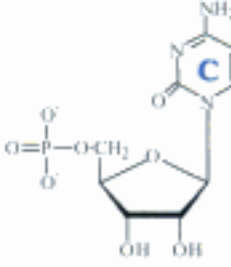
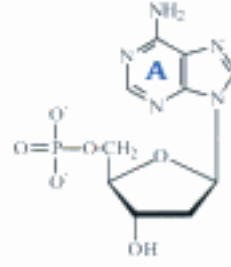
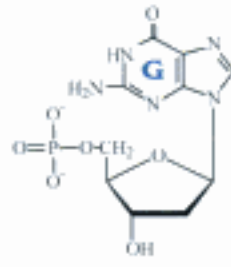
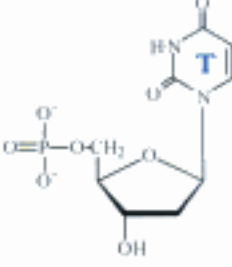
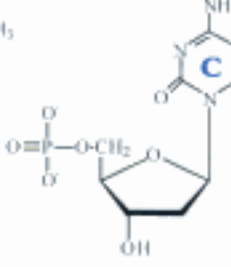
* (la desoxirribosa no tiene -OH en C2)

NOMENCLATURA DE NUCLEÓTIDOS SENCILLOS PIRIMIDÍNICOS

Bases:	Uracilo	Timina	Citosina	
	Ribonucleótidos	Desoxirribonucleótidos	Ribonucleótidos	Desoxirribo-nucleótidos
mono-fosfato	2'-UMP 2'-uridilato uridina-2'-fosfato 2'-uridina monofosfato	--- (*)	2'-CMP 2'-citidilato citidina-2'-fosfato 2'-citidina monofosfato	--- (*)
	3'-UMP (Up) 3'-uridilato uridina-3'-fosfato 3'-uridina monofosfato	3'-dTTP (Tp) 3'-desoxitimidilato desoxitimidina-3'-fosfato 3'-desoxitimidina monofosfato	3'-CMP (Cp) 3'-citidilato citidina-3'-fosfato 3'-citidina monofosfato	3'-dCMP desoxi-etc.
	5'-UMP (pU) 5'-uridilato uridina-5'-fosfato 5'-uridina monofosfato	5'-dTTP (pT) 5'-desoxitimidilato desoxitimidina-5'-fosfato 5'-desoxitimidina monofosfato	5'-CMP (pC) 5'-citidilato citidina-5'-fosfato 5'-citidina monofosfato	5'-dCMP desoxi-etc.
di-fosfato	5'-UDP (ppU) uridina-5'-difosfato 5'-uridina difosfato	5'-dTDP (ppT) desoxitimidina-5'-difosfato 5'-desoxitimidina difosfato	5'-CDP (ppC) citidina-5'-difosfato 5'-citidina difosfato	5'-dCDP desoxi-etc.
tri-fosfato	5'-UTP (pppU) uridina-5'-trifosfato 5'-uridina trifosfato	5'-dTTP (pppT) desoxitimidina-5'-trifosfato 5'-desoxitimidina trifosfato	5'-CTP (pppC) citidina-5'-trifosfato 5'-citidina trifosfato	5'-dCTP desoxi-etc.

* (la desoxirribosa no tiene -OH en C2)

3.2.2.2 Estructura de los nucleósidos-monofosfato que forman parte de los ácidos nucleicos

Nucleósidos-monofosfato componentes del RNA (= ribonucleótidos)			
Purínicos		Pirimidínicos	
 adenilato, AMP, 5'-AMP, A, adenosina-5'-fosfato	 guanilato, GMP, 5'-GMP, G, guanosina-5'-fosfato	 uridilato, UMP, 5'-UMP, U, uridina-5'-fosfato	 citidilato, CMP, 5'-UMP, C, citidina-5'-fosfato
Nucleótidos monofosfato componentes del DNA (= desoxirribonucleótidos)			
Purínicos		Pirimidínicos	
 desoxiadenilato, dAMP, 5'-dAMP, dA, desoxiadenosina-5'-fosfato	 desoxiguanilato, dGMP, 5'-dGMP, dG, desoxiguanosina-5'-fosfato	 desoxitimilato, dTTP, 5'-dTTP, dT, desoxitimidina-5'-fosfato	 desoxicitidilato, dCMP, 5'-dCMP, dC, desoxicitidina-5'-fosfato

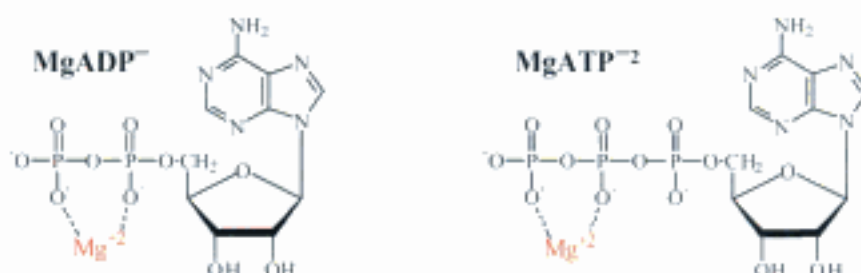
3.2.3 Propiedades fisicoquímicas de nucleósidos y nucleótidos

3.2.3.1 Propiedades iónicas

Como consecuencia de la disociación en disolución acuosa de los grupos **P**, **PP** y **PPP**, todos los nucleótidos (NMPs, NDPs y NTPs) muestran un marcado carácter ácido, con 2, 3 y 4 valores de pK_a , respectivamente. En general, al ser sus valores de pK_a inferiores a 7, se encuentran cargados negativamente (existen como aniones) a pH fisiológico. Gracias a este carácter aniónico, los nucleótidos forman **sales** con metales, son **crystalizables** y son más **solubles** en agua que los nucleósidos o sus bases nitrogenadas.

3.2.3.2 Formación de complejos

Los grupos **PP** de NDPs y **PPP** de NTPs muestran gran afinidad (NTP 10 veces más que NDP) por cationes divalentes, como Mg^{2+} , Mn^{2+} y Ca^{2+} . Al estar éstos presentes a alta concentración en la célula, los NDPs y NTPs no existen como aniones libres, sino formando **complejos quelato** 1:1, principalmente con Mg^{2+} . Bajo estas "formas activas" ($MgADP^{1-}$ y $MgATP^{2-}$) ambos nucleótidos participan en reacciones enzimáticas de transferencia de fosfato.



3.2.3.3 Absorción de la luz en el ultravioleta

Los nucleósidos y los nucleótidos muestran el característico máximo de **absorción** de la luz en el UV a 260 nm, debido a la base nitrogenada correspondiente. Por ejemplo, los valores de **absortividad molar** (ϵ , en $M^{-1} \text{ cm}^{-1}$) de algunos nucleósidos-monofosfato son los siguientes:

AMP	GMP	UMP	TMP	CMP
15.400	11.700	9.900	9.200	7.500

Veamos un **ejemplo de aplicación práctica** de esta propiedad:

Recuérdese que estos valores sirven para calcular la concentración, a partir de la fórmula

$$A_{260} = \epsilon_{260} \cdot b \cdot C \longrightarrow C = A_{260} / (\epsilon_{260} \cdot b)$$

ϵ = absortividad molar (antes llamada coeficiente de extinción molar)

b = paso óptico

C = concentración molar

Una mezcla de nucleótidos, obtenida por hidrólisis de una muestra de DNA, posee una absorbancia a 260 nm de 0,547 (medida en una cubeta de 1 cm de paso óptico). Suponiendo que la composición de ese DNA sea equimolar en los cuatro desoxirribonucleótidos, calcular la concentración milimolar de nucleótidos en la mezcla analizada.

Al tratarse de una mezcla equimolar, el coeficiente de extinción será la media:

$$\epsilon = (15.400 + 11.700 + 9.200 + 7.500) / 4 = 10.950 \text{ M}^{-1} \text{ cm}^{-1}$$

$$\text{luego: } C = (0,547 / 10.950) \text{ M} = 0,05 \text{ mM}$$

3.2.3.4 Nucleótidos como moléculas de reserva energética

Muchos nucleótidos (fundamentalmente, NDPs y NTPs) actúan en todo tipo de células y en multitud de procesos metabólicos como **compuestos ricos en energía**. Entre ellos, los nucleósidos trifosfato intervienen como precursores de la **síntesis de DNA y RNA**, con liberación de pirofosfato.

3.2.3.5 Hidrólisis

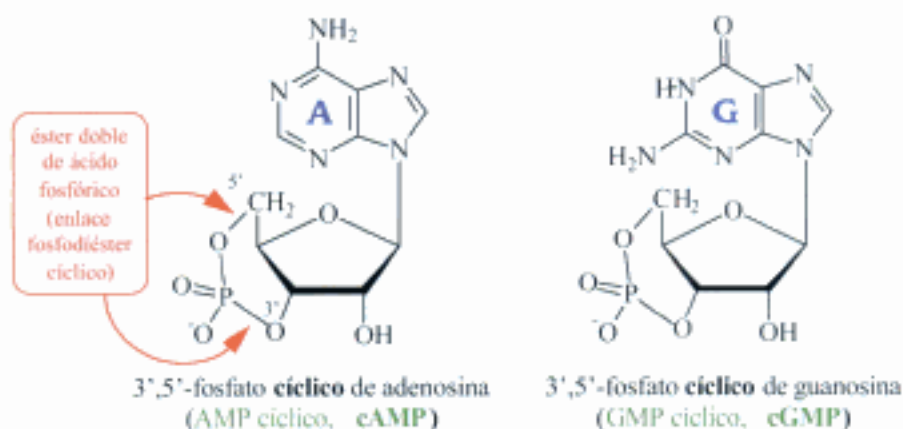
Los nucleósidos y nucleótidos pueden ser hidrolizados en sus componentes, en medio ácido o alcalino. Esta hidrólisis química posee cierto interés en el estudio de los ácidos nucleicos.



3.2.4 Nucleótidos cíclicos

Son nucleósidos-monofosfato en los que un mismo fosfato se une a dos posiciones del azúcar (dos -OH diferentes), generalmente 3' y 5'.

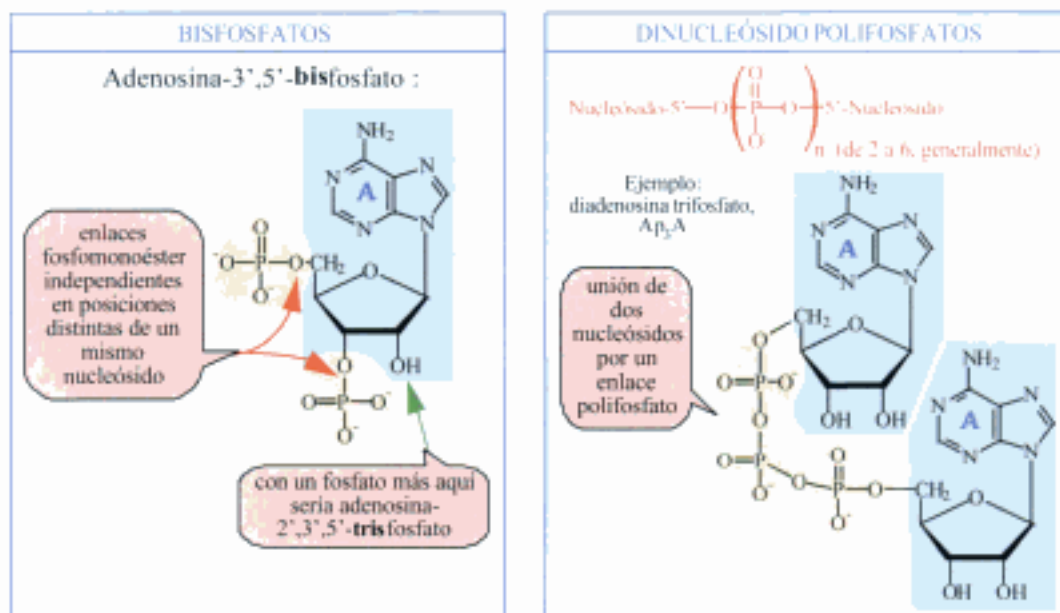
Actúan como segundos mensajeros, transmitiendo señales desde el entorno (por ejemplo, hormonas o primeros mensajeros) al interior celular.



3.2.5 Nucleótidos complejos

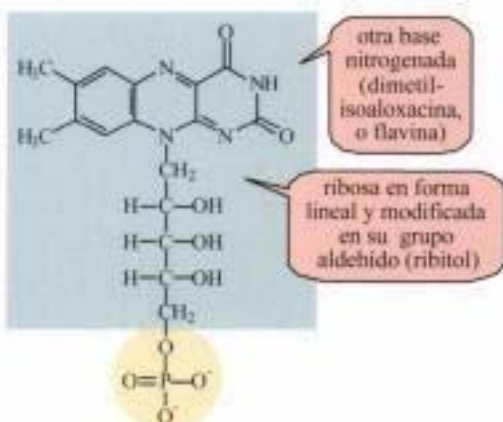
Finalmente, se engloban en este apartado aquellos nucleótidos, no componentes de los ácidos nucleicos y con funciones diversas, que estructuralmente difieren de los nucleótidos sencillos en uno o varios de los siguientes aspectos:

- Varios grupos fosfato esterificados en posiciones múltiples del azúcar (bis- y tris-fosfatos, CoA, NADP⁺).
- Presencia de dos nucleósidos, unidos entre sí por un "puente" de grupos fosfato (FAD, NAD⁺, Ap_nA); éstos entran dentro de la categoría de dinucleótidos.
- Presencia de otras bases nitrogenadas (FMN, FAD, NAD⁺, NADP⁺).
- Presencia de azúcares distintos de la ribofuranosa (FMN, FAD).
- Grupos funcionales adicionales (CoA).

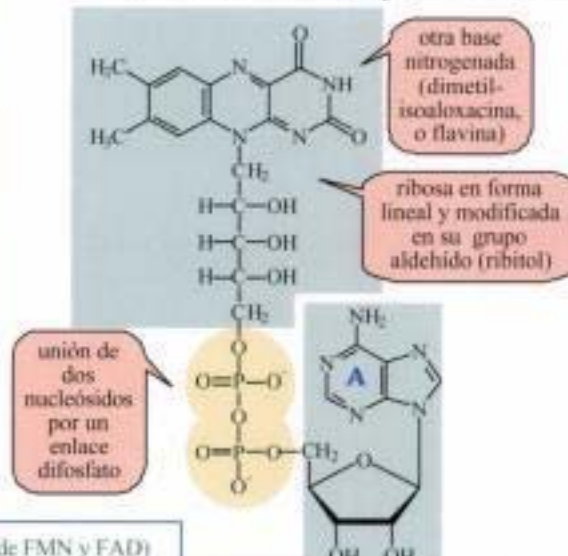
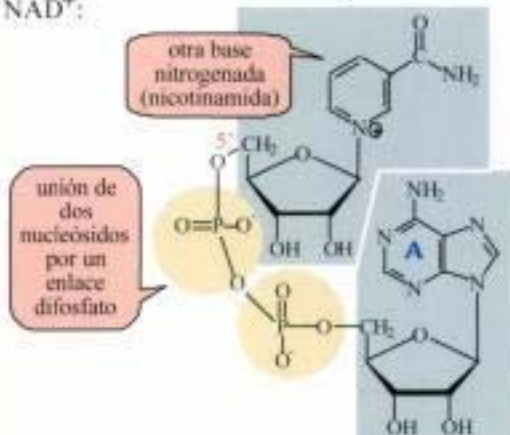
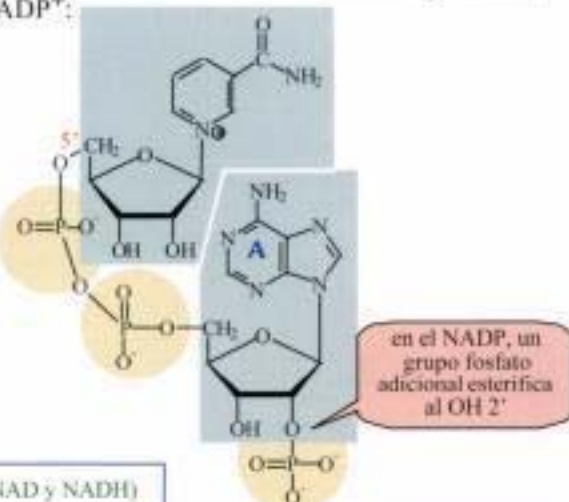


COENZIMAS REDOX

Mononucleótido de flavina, FMN:



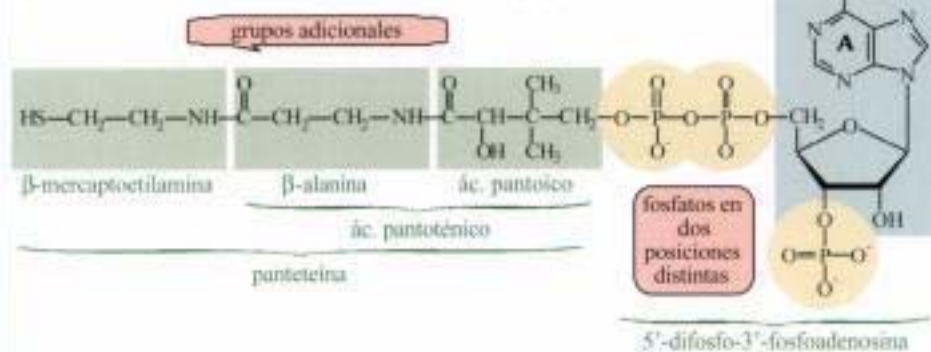
Dinucleótido de flavina y adenina, FAD:

(ribosa + flavina = riboflavina = Vitamina B₂, precursora de FMN y FAD)Dinucleótido de nicotinamida y adenina, NAD⁺:Fosfato del dinucleótido de nicotinamida y adenina, NADP⁺:

(nicotinato = niacina, vitamina precursora de los coenzimas NAD y NADH)

Coenzima A (CoA o CoA-SH):

(coenzima de acetilación; transportador universal de grupos acetilo)



Tema 4

Estructura primaria de ácidos nucleicos

4.1	Aspectos generales	31
4.1.1	Niveles estructurales	32
4.2	Estructura primaria	32
4.3	Formas de representación lineal	32
4.3.1	Fórmula completa	32
4.3.2	Representaciones esquemáticas	33
4.3.3	Representaciones abreviadas	33
4.4	Dos tipos de ácidos nucleicos según su composición	33
4.5	Propiedades fisicoquímicas de los ácidos nucleicos	34
4.5.1	Propiedades en disolución	34
4.5.2	Reactividad	34
4.5.3	Hidrólisis química de ácidos nucleicos	35
4.5.3.1	Hidrólisis ácida	35
4.5.3.2	Hidrólisis alcalina	35
4.5.4	Absorción en el ultravioleta	36

4.1 ASPECTOS GENERALES

Tradicionalmente se considera que los ácidos nucleicos forman parte de los **nucleoproteidos**, el grupo más importante de heteroproteidos o "proteínas conjugadas", entre los que también se incluyen glicoproteidos y lipoproteidos y, en ocasiones, también fosfoproteidos y cromo- o metaloproteidos. Efectivamente, los nucleoproteidos resultan de la interacción, mediante enlaces electrostáticos fuertes, de los residuos básicos (catiónicos) de ciertas proteínas (histonas, protaminas) y los grupos fosfato (aniónicos) del ácido nucleico. El estudio detenido de estas asociaciones nucleoproteicas se hará una vez conocida la estructura de los ácidos nucleicos a todos sus niveles.

Por su **composición química elemental** (C, H, O, N y P), derivada de la presencia exclusiva de nucleótidos en su estructura, los ácidos nucleicos se diferencian de las proteínas, esencialmente, en su mayor contenido de fósforo (10%) y en la ausencia total de azufre.

	Localización			Función
	Eucariotas	Procariontes	Virus	
DNA	en núcleo celular (varios cromosomas), en matriz mitocondrial y en estroma de cloroplastos	en la zona nucleocite del citosol (un cromosoma y varios plásmidos)	dentro de la cápsida (sólo en algunos tipos de virus)	Depositorio y transmisor de la información genética, organizada en genes que codifican productos génicos (proteínas o RNAs)
RNA	en núcleo celular (temporalmente), en citosol, en matriz mitocondrial y en estroma de cloroplastos	en citosol	dentro de la cápsida (en otros tipos de virus)	Interviene en la transmisión de la información desde el DNA hasta los productos génicos

Existen dos tipos de ácidos nucleicos, denominados **ribonucleico** o **RNA** y **desoxirribonucleico** o **DNA**, que difieren en los nucleótidos componentes. Antes de describir su estructura, recordemos brevemente que ambos ácidos nucleicos se encuentran en todo tipo de células, tanto procariotas como eucariotas; como excepción, los virus contienen sólo DNA o sólo RNA. En todos los casos, sus diversas funciones en la transmisión y expresión de la información genética vienen determinadas por su localización subcelular y su estructura de orden superior (conformación tridimensional).

4.1.1 Niveles estructurales

Desde un punto de vista estructural, en los ácidos nucleicos se pueden considerar **varios niveles**, similares en cierto modo a los descritos tradicionalmente para proteínas:

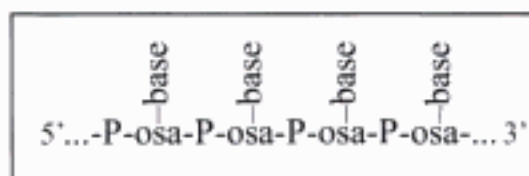
- **Estructura primaria**, descrita por la unión de numerosos nucleósidos mediante enlaces fosfodiéster, dando lugar a un polímero lineal. El orden de los nucleósidos (o sus bases) en la cadena define la **secuencia** del ácido nucleico.
- **Estructura secundaria**, con la que se inicia el análisis espacial de la molécula, correspondiente a la conformación local, es decir, la disposición relativa de nucleótidos que están próximos en la secuencia. Para el DNA este nivel estructural viene definido por la asociación de dos cadenas polinucleotídicas a través de las bases nitrogenadas que sobresalen del esqueleto azúcar-fosfato. En el RNA sólo se presenta en determinadas regiones de la molécula.
- **Estructuras de orden superior**: Se engloban bajo este término genérico todas aquellas estructuras tridimensionales que surgen a partir de los niveles primario y secundario. En eucariotas son complejas, variables y en ocasiones escasamente conocidas. En el caso del DNA pueden incluirse aquí las estructuras resultantes del superenrollamiento y de la asociación con proteínas básicas para formar la cromatina. Sin embargo, y a diferencia de lo que ocurre en proteínas, en general la estructura adoptada a este nivel no viene determinada por las de los niveles inferiores. En el caso de algunos RNAs (especialmente los tRNAs) se observa un plegamiento tridimensional bien definido, equiparable a la estructura terciaria de las proteínas. Aun en este caso, no se observa una estructura cuaternaria, ya que dicho plegamiento sólo afecta a una molécula y no a varias asociadas entre sí, como en las proteínas.

4.2 ESTRUCTURA PRIMARIA

La estructura primaria, común para DNAs y RNAs, consiste en la presencia de cadenas, generalmente largas, de carácter macromolecular, formadas por la unión de nucleósidos mediante **enlaces covalentes 3'-5'-fosfodiéster**. Por tanto, la unidad monomérica es el nucleótido, o nucleósido-monofosfato (generalmente contemplado como 5'-NMP).

Los ácidos nucleicos son, por tanto, polímeros de nucleótidos o **polinucleótidos**, con un **esqueleto lineal** formado por unidades alternas de fosfato y pentosa (parte constante de la secuencia), del cual sobresalen lateralmente las bases (parte variable, que define la secuencia), unidas al azúcar respectivo por enlaces N-glicosídicos.

Por convenio, la secuencia siempre se indica **en dirección 5' → 3'**, es decir, con el extremo 5' terminal (5'-P) a la izquierda y el extremo 3' terminal (3'-OH) a la derecha.



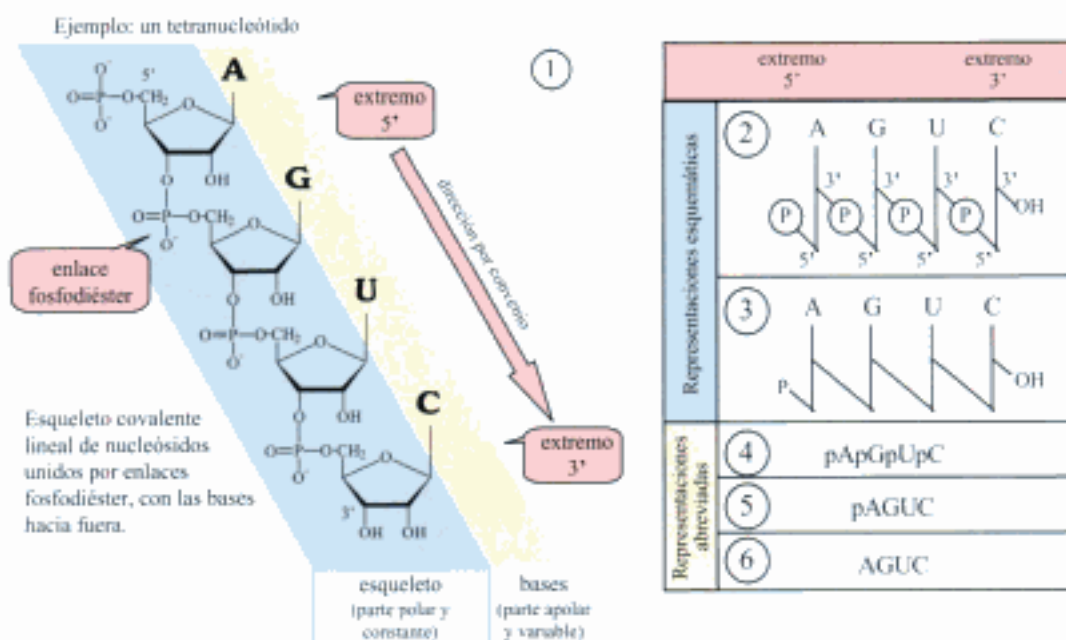
Dentro de este nivel estructural se estudiarán a continuación las distintas formas de representación, las diferencias en composición química de los dos tipos de ácidos nucleicos y sus propiedades fisicoquímicas más representativas.

4.3 FORMAS DE REPRESENTACIÓN LINEAL

Pueden emplearse hasta 6 niveles, de simplificación creciente. En general, se hace uso de la letra inicial de las bases, en mayúsculas, para representar a éstas o a sus nucleósidos. El ácido nucleico propuesto como ejemplo es un RNA (azúcar: ribosa y bases: A, G, U y C), pero las representaciones son igualmente válidas para un DNA (azúcar: desoxirribosa y bases: A, G, T y C).

4.3.1 Fórmula completa

Por razones de espacio y fundamentalmente didácticas, es aconsejable comenzar con una representación **vertical** del esqueleto del ácido nucleico (notación 1 en la figura), que permite apreciar con detalle los enlaces fosfodiéster entre nucleósidos consecutivos (uniones azúcar-fosfato) y la situación perpendicular al esqueleto de los enlaces N-glicosídicos (uniones azúcar-base). Para mantener la orientación 5' → 3', se sitúa el extremo 5'-P en la parte superior y el extremo 3'-OH en la inferior.



4.3.2 Representaciones esquemáticas

Como peculiaridad de esta representación, el anillo furanósico del azúcar se simplifica mediante un solo trazo vertical, cuyo extremo superior corresponde al C-1' y el inferior al C-5' (notación 2 en la figura).

En una forma aún más simple se omiten los fosfatos intermedios y la numeración de los azúcares (notación 3).

4.3.3 Representaciones abreviadas

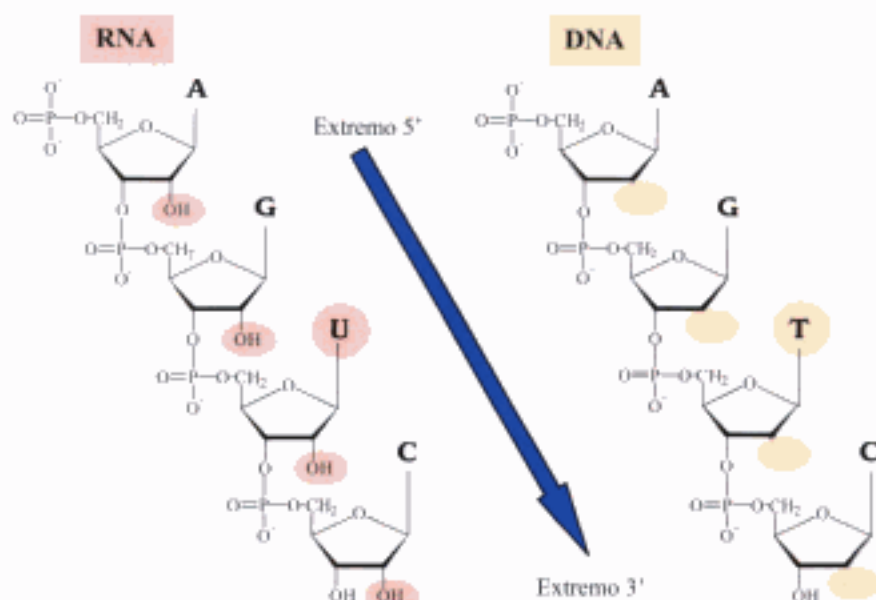
La simplificación se puede aumentar indicando cada nucleósido mediante la letra inicial de su base nitrogenada (lo único variable en la secuencia). El polinucleótido queda así reducido a una sucesión de letras mayúsculas intercaladas con letras **p** minúsculas (notación 4). El **extremo 5'-P** (a la izquierda) viene representado por una letra **p** seguida por la letra del primer nucleósido. El **extremo 3'-OH**, a la derecha, viene dado por la letra del último nucleósido. Los nucleósidos **intermedios** llevan la **p** a izquierda y derecha, correspondientes a las posiciones 5' y 3', respectivamente.

En la descripción cotidiana de secuencias de ácidos nucleicos se suprimen las letras **p** intermedias, con lo que la representación se reduce a una secuencia de letras precedida por una letra **p** para el fosfato en el extremo 5' (notación 5), o se asume la presencia del fosfato en 5' y se omite incluso esa primera **p** (notación 6).

4.4 DOS TIPOS DE ÁCIDOS NUCLEICOS SEGÚN SU COMPOSICIÓN

Tipo de ácido nucleico:		Ácido ribonucleico (RNA)	Ácido desoxirribonucleico (DNA)
Monómeros:		Ribonucleótidos	Desoxirribonucleótidos
Componentes del monómero	Fosfato	enlazando los monómeros (enlace fosfodiéster)	
	Azúcar	Ribosa	2'-Desoxirribosa
	Purinas	Adenina, A (6-aminopurina) Guanina, G (2-amino-6-oxopurina)	
	Pirimidinas	Citosina, C (2-oxo-4-aminopirimidina) Uracilo, U (2,4-dioxopirimidina)	
		Uracilo, U (2,4-dioxopirimidina)	Timina, T (5-metiluracilo)

Según la notación de fórmula completa, la representación de ambos ácidos nucleicos es la siguiente:



4.5 PROPIEDADES FISICOQUÍMICAS DE LOS ÁCIDOS NUCLEICOS

Los ácidos nucleicos presentan propiedades importantes en relación con sus funciones de almacenamiento, transmisión (replicación) y expresión (transcripción y traducción) de la información genética. Se comentan aquí brevemente algunas de ellas, relacionadas con la estructura, y el resto se irán estudiando a medida que surja su aplicación al propio estudio de estas moléculas (por ejemplo, la renaturalización de un DNA desnaturizado se analizará como fundamento de la hibridación).

4.5.1 Propiedades en disolución

Las cadenas polinucleotídicas de DNA y RNA son **hidrofilicas**, debido a la posibilidad de formación de enlaces de hidrógeno con el agua por parte de los numerosos grupos fosfato y -OH libres del azúcar a lo largo del esqueleto.

Por otro lado, a pH fisiológico los grupos fosfato (pK_a próximo a 6) se ionizan casi por completo, haciendo que las moléculas de DNA y RNA se comporten como **ácidos** y tengan numerosas cargas negativas. Debido a este carácter **polianiónico**, los ácidos nucleicos se encuentran casi siempre neutralizados por interacción iónica con las cargas positivas de otras moléculas. Concretamente, el DNA del genoma nuclear eucariótico se encuentra unido a **histonas**, proteínas ricas en Arg y Lys, cargadas positivamente a pH fisiológico, dando lugar a nucleoproteidos; la importancia de esta asociación se considerará posteriormente (pág. 84). En los espermatozoides, el papel de las histonas es desempeñado por pequeñas proteínas ricas en Arg, llamadas **protaminas** y también con fuerte carga positiva. En otras ocasiones, el DNA se une a **poliaminas** (principalmente espermina y espermidina, así llamadas por haber sido detectadas originalmente en el semen), distribuidas ampliamente pero de forma especial en algunos virus y bacterias en estado de rápida proliferación. Así, por ejemplo, en el fago T₄ las poliaminas neutralizan un 40% de la carga negativa total del DNA, estabilizando su conformación. Por último, como resultado experimental de esta propiedad, la cristalización de ácidos nucleicos se facilita si su carga negativa se compensa por la unión de iones metálicos.

Debido a la relativa rigidez de la molécula (en doble hélice) y a su inmensa longitud en relación al diámetro, las soluciones de DNA son muy **viscosas**. Esta propiedad tiene interés en relación con el estudio del proceso de desnaturización (pág. 165).

4.5.2 Reactividad

En general, los ácidos nucleicos son químicamente muy estables, especialmente en el caso del DNA. Ello se debe a que la desoxirribosa carece de grupos reactivos: el 4'-OH está comprometido en el cierre del anillo, y los 3'-OH y 5'-OH están formando enlaces fosfodiéster. El RNA es algo más reactivo, gracias a los grupos 2'-OH, lo que se refleja en sus propiedades, por ejemplo, su hidrólisis en medio alcalino, como se verá a continuación.

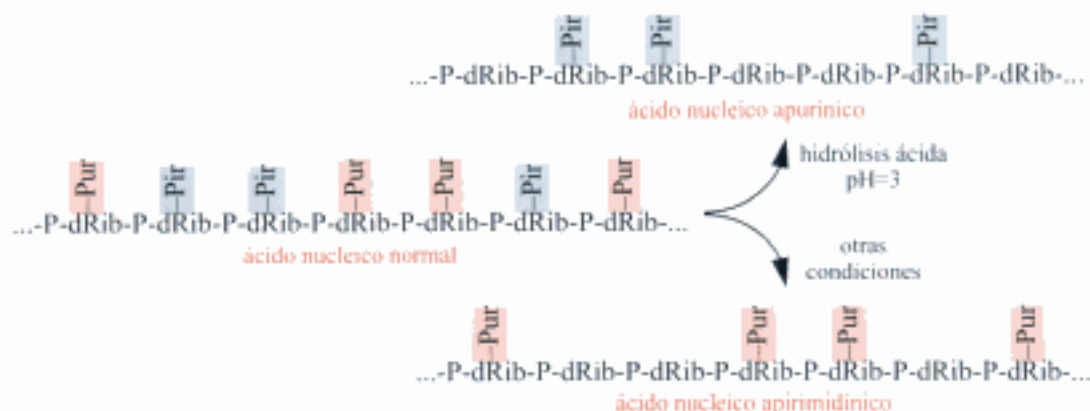
4.5.3 Hidrólisis química de ácidos nucleicos

4.5.3.1 Hidrólisis ácida

Los **ácidos fuertes** escinden tanto los enlaces fosfodiéster entre nucleósidos como los enlaces N-glicosídicos de cada nucleótido. Por tanto, sirven para analizar la composición de bases de un ácido nucleico, pero no para determinar su secuencia.

¿se hidrolizan?	DNA		RNA	
	Enlaces fosfodiéster (entre nucleósidos)	Enlaces N-glicosídicos (en cada nucleósido)	Enlaces fosfodiéster (entre nucleósidos)	Enlaces N-glicosídicos (en cada nucleósido)
Medio ácido fuerte (p.ej., HCl 1M)	sí	sí	sí	sí
Medio ácido débil	no	sí (todos o purinas)	no	sí (todos o purinas)

Los **ácidos débiles** respetan los enlaces fosfodiéster, pero tienden a romper la unión N-glicosídica. Este enlace es más lábil con purinas que con pirimidinas, por lo que se puede conseguir (por ejemplo, a pH=3) escindir selectivamente las bases purínicas, dando lugar a los denominados ácidos *apurínicos*. Esta hidrólisis es de cierto interés para estudiar parcialmente la secuencia del DNA. Con el mismo fin, también se han preparado ácidos nucleicos *apirimidínicos*, por eliminación selectiva de las pirimidinas bajo condiciones químicas diferentes.

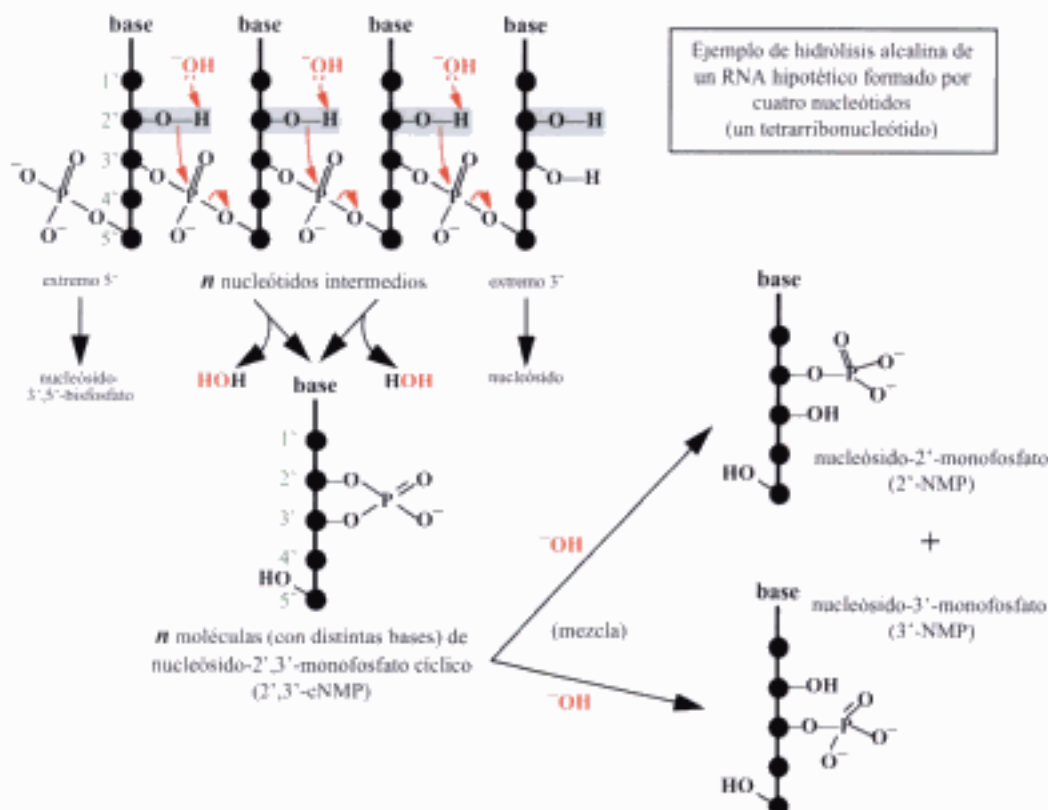


4.5.3.2 Hidrólisis alcalina

El medio básico afecta de distinta manera a los enlaces fosfodiéster, hidrolizando los del RNA pero no los del DNA; los enlaces N-glicosídicos son resistentes.

¿se hidrolizan?	DNA		RNA	
	Enlaces fosfodiéster (entre nucleósidos)	Enlaces N-glicosídicos (en cada nucleósido)	Enlaces fosfodiéster (entre nucleósidos)	Enlaces N-glicosídicos (en cada nucleósido)
Medio alcalino	no	no	sí	no

El **RNA** se hidroliza rápidamente, liberando sus nucleótidos, en disoluciones diluidas de NaOH (pH 11), sin afectar a los enlaces N-glicosídicos. La hidrólisis de los enlaces azúcar-fosfato se debe a la implicación directa de los grupos 2'-OH de la ribosa en la catálisis química. Se forman así NMPs 2',3'-cíclicos como productos intermedios, que son hidrolizados rápidamente en el medio alcalino, al azar a uno u otro lado del enlace 2',3'-fosfodiéster, para dar una mezcla de los 2'-NMPs y 3'-NMPs.



El **DNA**, al contrario que el RNA, es **resistente a la hidrólisis** en medio alcalino. Ello se debe a la ausencia de grupos 2'-OH en las unidades de desoxirribosa, lo que impide el mecanismo expuesto anteriormente.

4.5.4 Absorción en el ultravioleta

La similitud en el valor máximo de absorción de las distintas bases (pág. 17) puede utilizarse no sólo para detectar y cuantificar las bases, los nucleósidos y los nucleótidos, sino especialmente para determinar la concentración de los ácidos nucleicos. De hecho, es el procedimiento comúnmente utilizado, a pesar de no ser el más exacto. Basta con medir el valor de A_{260} en una disolución diluida, suficientemente transparente, de DNA o RNA para poder calcular su concentración. Sin embargo, es conveniente la realización de algunos controles, teniendo en cuenta especialmente la posible contaminación con proteínas, cuya presencia contribuye a la absorción y, por tanto, interfiere en esta determinación.

Efectivamente, aunque los máximos de absorción son característicos de cada uno de estos compuestos (260 nm para ácidos nucleicos y 280 nm para proteínas), ambos absorben en cierta medida a las dos longitudes de onda. Por ejemplo, la absorbancia de proteínas a 260 nm es, para igual concentración, unas 20 ó 30 veces inferior a la de los ácidos nucleicos. En la práctica, teniendo en cuenta esta situación, se deben emplear ambos valores de absorción (A_{260} y A_{280}) para poder corregir la interferencia mutua y así determinar correctamente las concentraciones de ácido nucleico y de proteína.

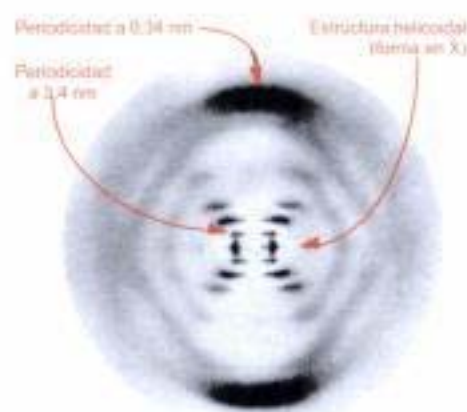
Tema 5

Estructura secundaria del B-DNA

5.1	Proporción de bases nitrogenadas: Reglas de Chargaff.....	38
5.1.1	Algunas relaciones entre bases son iguales en todas las especies.....	38
5.1.1.1	Relación entre bases.....	38
5.1.1.2	Relación entre purinas y pirimidinas.....	38
5.1.1.3	Relación entre aminobases y oxobases.....	39
5.1.2	Otras relaciones son diferentes según la especie.....	39
5.2	Modelo de Watson y Crick: forma B del DNA.....	39
5.2.1	Complementariedad de las bases nitrogenadas.....	39
5.2.1.1	Consecuencias de la complementariedad.....	41
5.2.1.2	Efecto del pH sobre el apareamiento de bases.....	42
5.2.2	Geometría de los pares de bases.....	43
5.2.3	Antiparalelismo de las dos hebras.....	44
5.2.4	Diferencias en la secuencia de ambas hebras.....	44
5.2.5	Helicidad y carácter dextrógiro.....	45
5.2.6	Parámetros cuantitativos de la doble hélice.....	46
5.2.7	Carácter anfipático y estabilidad de la doble hélice.....	47
5.2.7.1	Situación de los fosfatos y carácter hidrofílico.....	47
5.2.7.2	Coplanariedad, apilamiento de las bases y carácter hidrofóbico.....	48
5.2.8	Superficie de la molécula: surcos en el DNA.....	48
5.2.9	Significado biológico: idoneidad para la replicación.....	49

El análisis por difracción de rayos X ha influido notablemente sobre el conocimiento de la estructura secundaria del DNA. Desde sus inicios se observaron, incluso en muestras impuras de DNA disueltas, señales que correspondían a un espaciado regular de 0,34 nm a lo largo de las fibras de DNA. En los años 50 se obtuvieron datos más precisos, con muestras de DNA más puro y haciendo uso de los métodos de difracción de rayos X, que ya se habían desarrollado para el estudio estructural de proteínas. A partir de los correspondientes patrones de difracción, donde las manchas centrales en forma de aspa indican una estructura en hélice y las bandas intensas en la parte superior e inferior indican la repetición periódica de estructuras, se concluyó que la molécula de DNA estaba formada por dos cadenas, con una disposición helicoidal y con una doble periodicidad a lo largo del eje de la hélice, cada 0,34 nm y cada 3,4 nm.

Por otra parte, la química de nucleótidos permitía conocer la existencia de interacciones hidrofóbicas por apilamiento, así como la asociación mediante puentes de hidrógeno ("apareamiento"). Todos estos datos experimentales, junto con el análisis de la proporción de las bases constituyentes de la molécula, cuyas conclusiones se exponen a continuación, y la construcción de modelos moleculares, sirvieron de base para la propuesta por Watson y Crick de la estructura de doble hélice, cuya descripción es el objeto principal de este tema.



5.1 PROPORCIÓN DE BASES NITROGENADAS: REGLAS DE CHARGAFF

A pesar de que la composición de bases puede parecer un aspecto ligado a la estructura primaria, su estudio se hace en este tema por estar directamente relacionada con la estructura secundaria del DNA. En un principio se creyó que el DNA estaba formado por cantidades iguales de los cuatro nucleótidos A, G, C y T, y que la unidad fundamental era el tetranucleótido, hasta que Erwin Chargaff, entre otros, observó distintas proporciones de los 4 nucleótidos en diversos organismos. Concretamente, los DNAs aislados a partir de células o tejidos diferentes de distintos individuos, dentro de una misma especie, presentan idéntica composición de bases, que además **no cambia** con la edad, estado nutricional o variaciones del entorno. La composición del DNA es, por tanto, característica de cada especie. A partir de esos datos se formularon unas **reglas o generalizaciones cuantitativas sobre la composición de bases del DNA**, que fueron la clave para deducir su estructura en doble hebra.

	Composición de bases (% moles)				Relación entre bases			
	A Adenina	T Timina	G Guanina	C Citosina	$\frac{A}{T}$	$\frac{G}{C}$	Purinas/ Pirimidinas $\frac{A+G}{T+C}$	Relación de asimetría $\frac{A+T}{G+C}$
Animales:								
Hombre	30,9	29,4	19,9	19,8	1,05	1,01	1,03	1,52
Oveja	29,3	28,3	21,4	21,0	1,04	1,02	1,03	1,36
Rata	28,6	28,4	21,4	20,4	1,01	1,05	1,02	1,36
Gallina	28,8	29,2	20,5	21,5	0,99	0,95	0,97	1,38
Tortuga	29,7	27,9	22,0	21,3	1,06	1,03	1,05	1,33
Salmón	29,7	29,1	20,8	20,4	1,02	1,02	1,02	1,43
Erizo de mar	32,8	32,1	17,7	17,3	1,02	1,02	1,02	1,85
Bacterias:								
Escherichia coli	24,7	23,6	26,0	25,6	1,05	1,02	1,03	0,94
Staphylococcus aureus	30,8	29,2	21,0	19,0	1,05	1,11	1,07	1,50
Chlostridium perfringens	36,9	36,3	14,0	12,8	1,02	1,09	1,04	2,73
Brucella abortus	21,0	21,1	29,0	28,9	1,00	1,00	1,00	0,73
Sarcina lutea	13,4	12,4	37,1	37,1	1,08	1,00	1,02	0,35
Virus:								
Fago ϕ 174 (monofilar)	24,3	32,7	24,1	18,5	0,74	1,30	0,95	1,34
Fago ϕ 174, forma replicativa (bifilar)	26,3	26,4	22,3	22,3	1,00	1,00	1,00	1,18

5.1.1 Algunas relaciones entre bases son iguales en todas las especies

A pesar de que la composición varía entre especies, en todas ellas se cumplen ciertas relaciones entre las bases, formuladas como las tres primeras reglas de Chargaff, que se exponen a continuación.

5.1.1.1 Relación entre bases

En prácticamente todas las especies, el contenido de A se aproxima al de T, y el contenido de G al de C:

$A \approx T$	$\rightarrow$	Razón	$A/T \approx T/A \approx 1$
$G \approx C$	$\rightarrow$	Razón	$G/C \approx C/G \approx 1$

Esta observación fue esencial para que Watson y Crick postulasen los pares AT y GC, que definen la complementariedad de bases en el DNA de doble hebra. Hay que destacar que esta regla no se cumple para las pocas especies cuyo material genético está formado por RNA o DNA de hebra sencilla (compárense en la tabla los ejemplos del virus ϕ 174, en sus formas de hebra sencilla y de doble hebra).

5.1.1.2 Relación entre purinas y pirimidinas

Como consecuencia de lo anterior, el contenido de bases purínicas (A + G) se aproxima al de bases pirimidínicas (T + C), es decir, hay un 50% de purinas y un 50% de pirimidinas:

$A+G \approx T+C$	$\approx 50\%$	$\rightarrow$	Razón	$(A+G) / (T+C) \approx 1$			
purinas	$\approx$	pirimidinas	$\approx 50\%$	$\rightarrow$	Razón	purinas / pirimidinas	≈ 1

5.1.1.3 Relación entre aminobases y oxobases

Al observarse que $A/T = C/G = 1$, también se ha de cumplir (por razones estrictamente matemáticas) que $(A+C) / (T+G) = 1$. Por tanto, puede concluirse, como relación adicional (sin un significado especial), que:

$$6\text{-aminopurinas (A)} + 4\text{-aminopirimidinas (C)} = 6\text{-oxopurinas (G)} + 4\text{-oxopirimidinas (T)}$$

5.1.2 Otras relaciones son diferentes según la especie

Otras relaciones entre bases no se mantienen constantes entre especies; de entre ellas, resulta útil la denominada **razón de asimetría**:

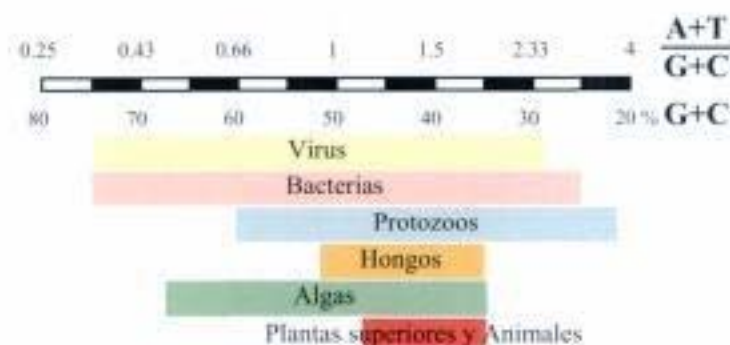
$$\frac{A + T}{G + C}$$

que, además de tener un valor diferente para cada especie, no se aproxima a la unidad como las razones anteriores (ver tabla). Por ejemplo, en animales la razón de asimetría es siempre superior a uno; lo mismo ocurre en plantas superiores y hongos. Por contraste, en bacterias y virus los valores son mucho más variados, superiores o inferiores a la unidad.

La razón de asimetría puede también calcularse directamente a partir del contenido en G+C, parámetro habitual en otros tipos de estudio, como la desnaturalización del DNA:

$$\text{razón de asimetría} = \frac{A + T}{G + C} = \frac{\%(A+T)}{\%(G+C)} = \frac{100 - \%(G+C)}{\%(G+C)}$$

De acuerdo con ello, los animales, plantas superiores y hongos muestran un exceso de A+T (50-65%) sobre G+C (50-35%), mientras que en bacterias y virus se aprecia la gran variabilidad en G+C (75-25%).



En resumen, con estos estudios se ha observado que los DNAs de especies evolutivamente relacionadas tienen una razón de asimetría similar, mientras que las especies muy alejadas tienden a tener una razón diferente. A pesar de ser ésta la menos conocida de las reglas de Chargaff, el hecho de que permita distinguir entre especies hace que pueda utilizarse como criterio en las relaciones evolutivas y las clasificaciones taxonómicas de los organismos.

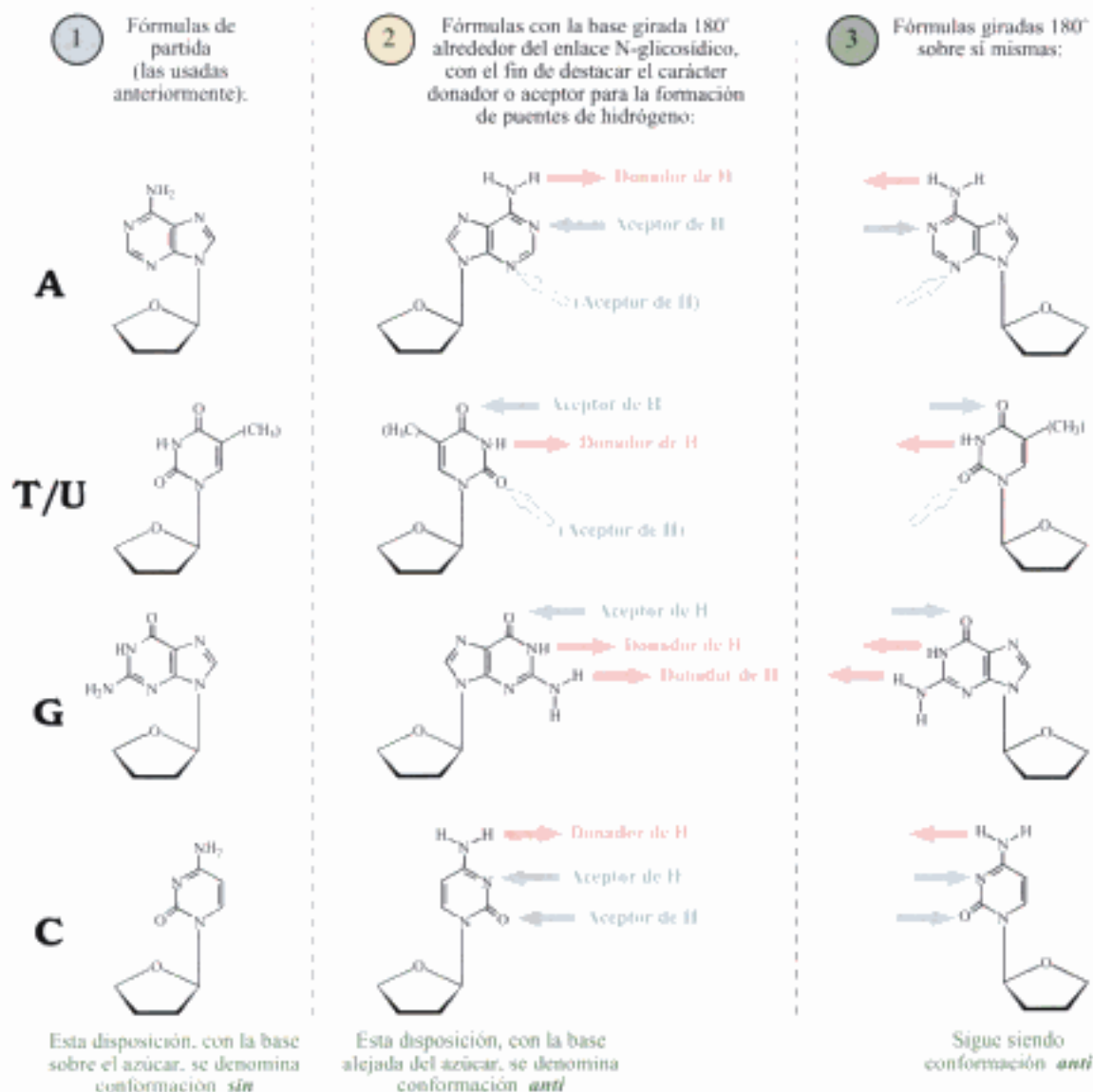
5.2 MODELO DE WATSON Y CRICK: FORMA B DEL DNA

El **modelo tridimensional**, como estructura de referencia para las propiedades del DNA, fue postulado por James Watson y Francis Crick en 1953. Corresponde a la actualmente denominada **forma B del DNA**, la más común en las células y la más estable de todas las que puede adoptar un DNA bajo condiciones fisiológicas. Este modelo se basó tanto en las observaciones por rayos X como en las reglas de Chargaff y, por tanto, está de acuerdo con ellos, así como con otros estudios realizados mediante un gran número y variedad de técnicas. Simultáneamente se propuso la adecuación del modelo a la funcionalidad de la molécula de DNA para la conservación y transmisión de la información genética (pág. 49). Presenta las siguientes características básicas:

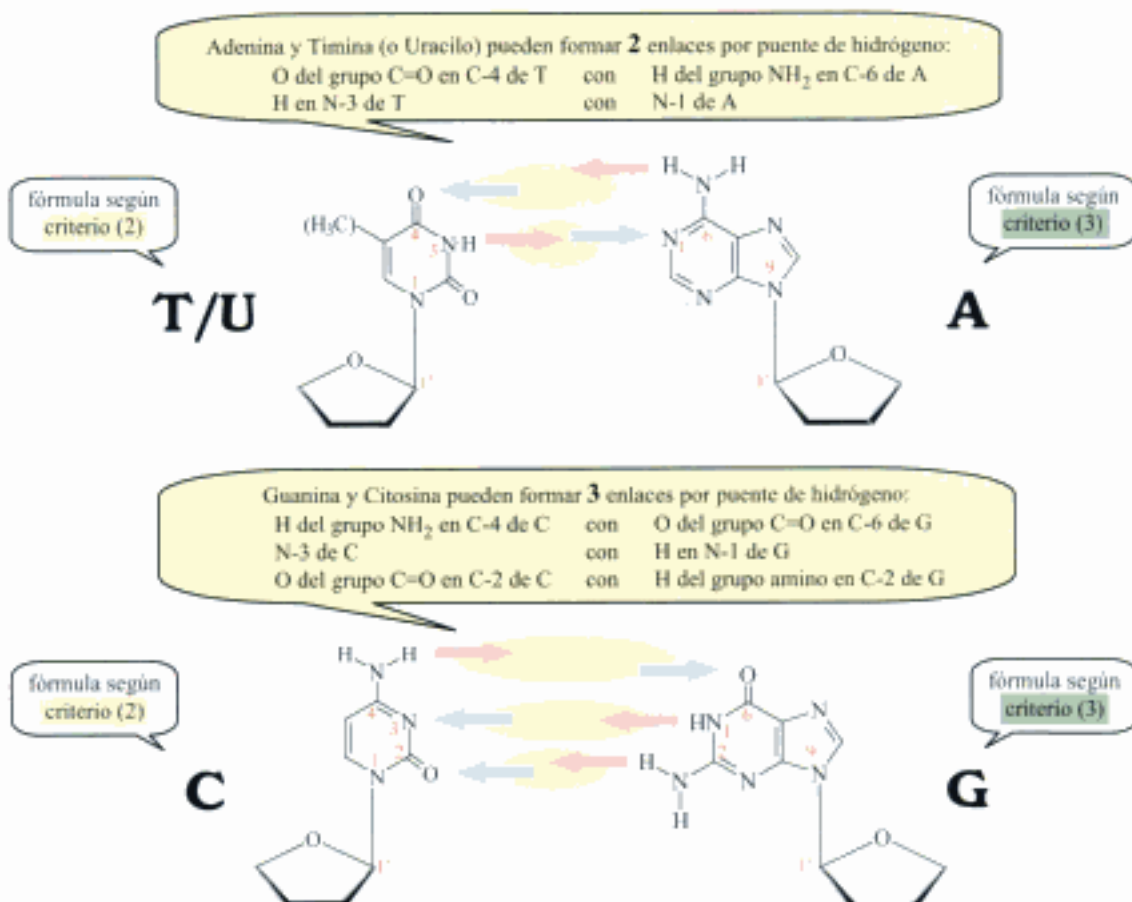
5.2.1 Complementariedad de las bases nitrogenadas

La estructura química de las bases permite la formación de enlaces por puente de hidrógeno (dos átomos electronegativos, como N y O, comparten un átomo de hidrógeno). En concreto, dos nucleótidos pueden interactuar de forma **específica** mediante dichos enlaces y se habla de **"bases complementarias"**, **"pares de bases"** o **"apareamiento de bases"**.

Para una mejor comprensión de la complementariedad, es conveniente establecer unos criterios previos en cuanto a la escritura de los nucleótidos en cada pareja de bases. Se trata con ello, simplemente, de facilitar la descripción de unas estructuras relativamente complejas:



- 4 Al representar un par de bases, se emplea la fórmula de tipo 2 para el nucleótido situado a la izquierda, y la de tipo 3 para el situado a la derecha (ambas conformación *anti*)



Las interacciones óptimas se forman entre G y C (tres enlaces de hidrógeno) y entre A y T, o bien A y U (dos enlaces de hidrógeno). El par G=C establece así una interacción más fuerte que el par A=T (en el DNA) o el par A=U (en el RNA). Obsérvese que siempre se forma un enlace de hidrógeno entre los dos átomos de N nucleares, mientras que los restantes se forman entre un grupo amino exocíclico y un grupo carbonilo.

Son también posibles otros apareamientos entre bases, denominados "de tipo Hoogsteen", que aparecen menos frecuentemente y no forman parte del B-DNA. Por ejemplo, en el RNA (pág. 73), en las triples hélices de DNA o RNA (pág. 60) y en la interacción codón-anticodón (pág. 296).

5.2.1.1 Consecuencias de la complementariedad

Gracias a la complementariedad descrita, dos cadenas polinucleotídicas se pueden asociar siempre y cuando todas sus bases sean complementarias. Este apareamiento completo de las dos cadenas da lugar a un ácido nucleico de **doble hebra** o **doble cadena**, estructura habitual del DNA. Esto quiere decir que siempre que en un DNA exista A, G, C o T en una cadena, se encontrarán T, C, G y A, respectivamente, en la otra.

Es importante destacar que la complementariedad entre las dos hebras del DNA es fundamental para su papel como portador de la información genética. Por un lado, porque se reduce la posibilidad de que las bases se modifiquen químicamente. Por otro, porque se facilita la corrección de mutaciones y de errores producidos durante la replicación del DNA. Finalmente, porque la información está en cierto modo duplicada, aunque las secuencias de ambas hebras son diferentes entre sí (pág. 44).

Además, la existencia de la complementariedad explica las reglas de Chargaff (pág. 38), y justifica dos características estructurales importantes, la helicidad y el antiparalelismo, que se estudian más adelante.

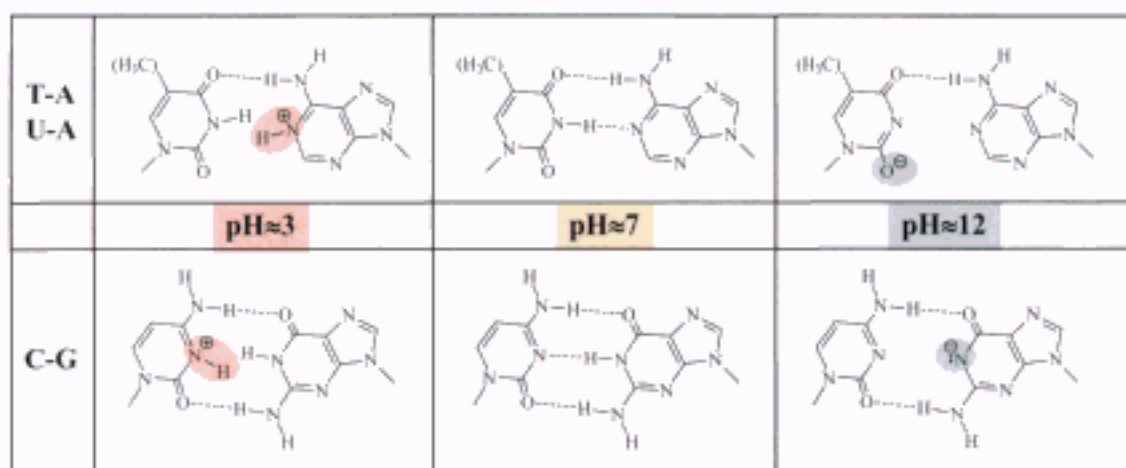
Ilustración, mediante un ejemplo hipotético basado en la composición del DNA humano, de cómo la doble hebra del DNA permite explicar las reglas de Chargaff:

bases	Composición de bases			Regla de Chargaff correspondiente:
	% moles en hebra 1 y sus cocientes	% moles en hebra 2 y sus cocientes	% moles en la doble hebra y sus cocientes	
A	32	28	30	Relación entre bases individuales: $A=T$, $C=G$
T	28	32	30	
G	25	15	20	
C	15	25	20	
A+G = purinas	57	43	50	Relación entre purinas y pirimidinas: $A+G / C+T = 1$
C+T = pirimidinas	43	57	50	
purinas / pirimidinas	1,33	0,75	1,0	
A+C = aminobases	47	53	50	Relación entre aminobases y oxobases $= 1$
G+T = oxobases	53	47	50	
amino / oxobases	1,08	0,92	1,0	

El apareamiento implica que el contenido de una base en una hebra ha de ser igual al de su base complementaria en la otra hebra. Por ello, las reglas de Chargaff se cumplen cuando se considera la **composición total del DNA de doble hebra**, pero no para **cada hebra por separado**, ya que en este caso, $A \neq T$ y $C \neq G$. Tampoco se cumplen las reglas en los ácidos nucleicos de cadena sencilla (p.ej., RNA y DNA vírico), donde $A \neq U$, $A \neq T$ y $C \neq G$.

5.2.1.2 Efecto del pH sobre el apareamiento de bases

Los cambios de pH, al afectar a la ionización de los grupos funcionales de las bases nitrogenadas y al equilibrio de tautómeros, modifican las posibilidades de formación de enlaces de hidrógeno y, por tanto, el apareamiento de bases.



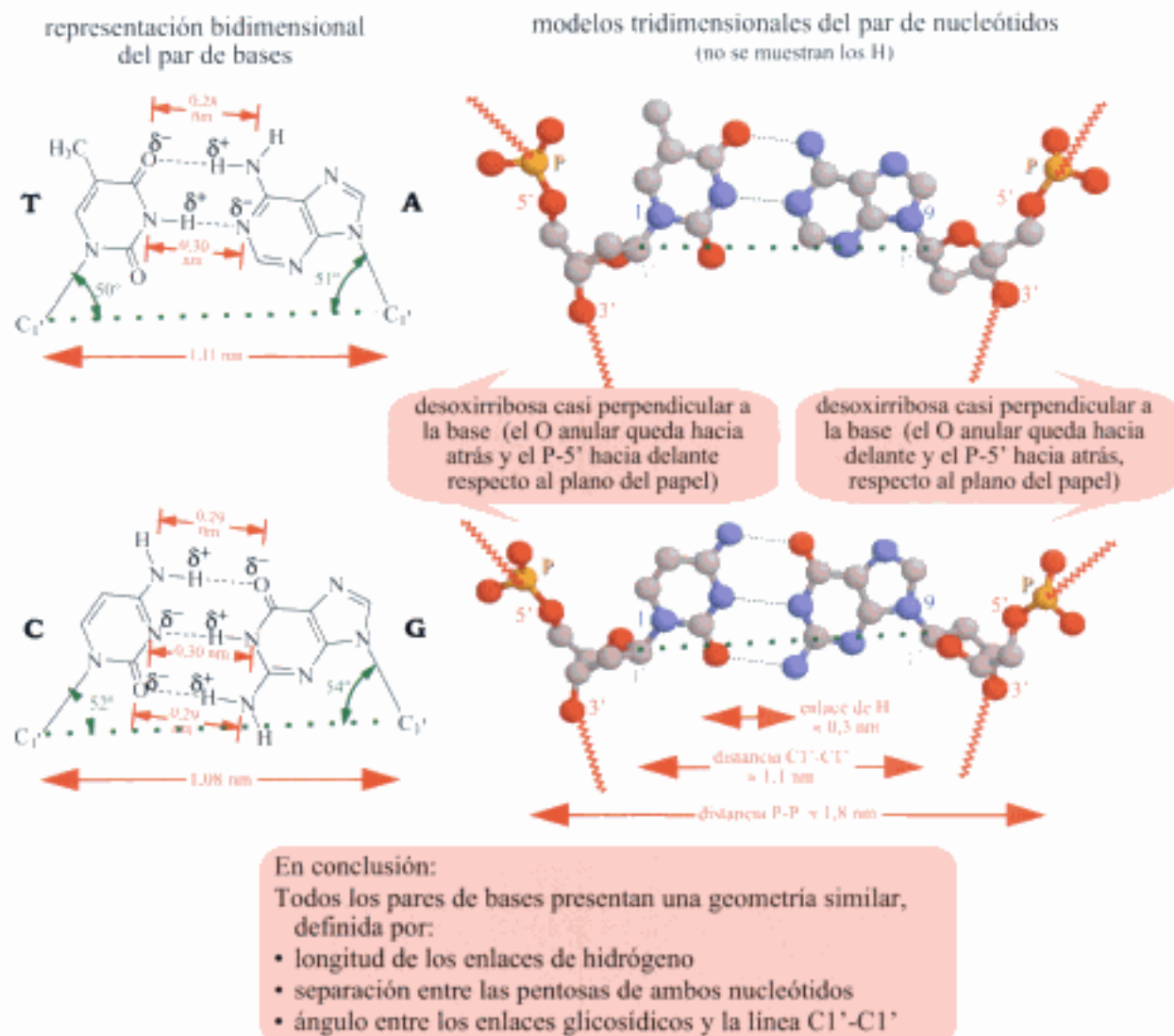
(cada base tiene otras formas resonantes y tautoméricas, no indicadas)

Como consecuencia, a valores de pH inferiores o superiores al fisiológico el menor número de enlaces de hidrógeno posibles provoca la separación de las hebras componentes del DNA, fenómeno denominado **desnaturalización** (pág. 164). No debe confundirse este efecto con la hidrólisis ácida o alcalina, bajo condiciones más extremas, de las cadenas de ácido nucleico (pág. 35).

5.2.2 Geometría de los pares de bases

A pesar de la dificultad de visualizar la geometría de un par de bases en ausencia de un modelo cercano a la realidad (modelo tridimensional, sea material o informático), se describe a continuación de la forma más fiel posible.

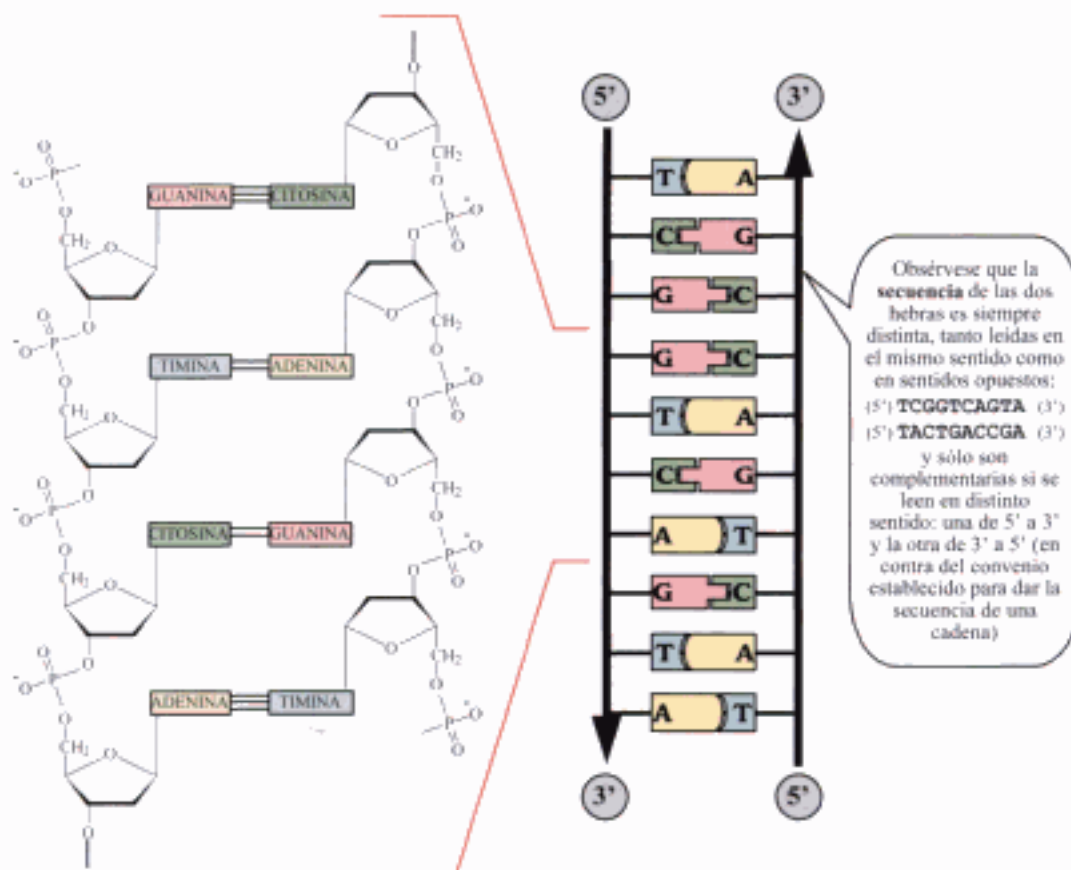
La formación de un enlace de hidrógeno requiere que los tres átomos implicados (dos electronegativos y un H) se sitúen sobre una línea recta. Esta condición sólo se puede cumplir cuando los dos enlaces N-glicosídicos de los nucleósidos complementarios forman un ángulo, inferior a 90° , con la línea imaginaria que une los C-1' de ambas pentosas. Esta geometría del par de bases determina la disposición espacial relativa de las dos hebras en el DNA (helicidad) y es la causa a nivel molecular de la presencia de surcos menor y mayor en su estructura tridimensional; ambos aspectos se analizan en detalle más adelante (págs. 45 y 48).



Una vez establecidos los enlaces de hidrógeno múltiples, las distancias entre los átomos que los forman son prácticamente iguales para todos los casos. Igualmente, al aparearse siempre una purina con una pirimidina, la separación entre las pentosas de ambas hebras también se mantiene constante. Sólo de esta manera (distancia constante entre las hebras, independientemente de su secuencia) es posible la asociación continua de las dos cadenas para formar una molécula de DNA doble.

5.2.3 Antiparalelismo de las dos hebras

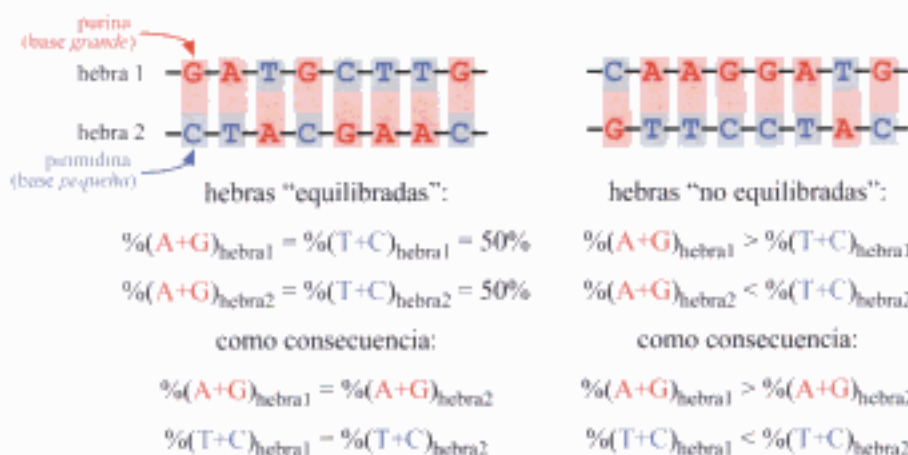
Ya se ha indicado que el apareamiento completo entre las dos hebras exige que sus secuencias sean complementarias. Debe añadirse ahora que estructuralmente esto sólo se puede conseguir si ambas hebras se disponen en sentidos opuestos: si una avanza de 5' a 3', la otra debe hacerlo de 3' a 5'. Se dice que las hebras tienen polaridad opuesta, o que son **antiparalelas**. El concepto de antiparalelismo va, pues, estrechamente ligado al de complementariedad.



5.2.4 Diferencias en la secuencia de ambas hebras

Debe destacarse que las dos hebras de un DNA, complementarias y antiparalelas, poseen secuencias totalmente diferentes, como se puede ver en el último ejemplo mostrado. Es decir, son **hebras no idénticas**. Este hecho, aparentemente simple pero que a menudo pasa inadvertido, es de gran trascendencia; en concreto, en el proceso de transcripción, la molécula de RNA que se sintetiza es completamente diferente según cuál de las dos hebras del DNA se transcriba.

En ocasiones se utiliza el término *hebras equilibradas* refiriéndose a aquellas regiones del DNA en las que una de las hebras contiene igual proporción de purinas y pirimidinas. Como consecuencia, la hebra complementaria es también equilibrada:



En principio, la existencia de hebras equilibradas o no se refiere a regiones localizadas de una molécula de DNA. Existe, sin embargo, un ejemplo singular en el DNA mitocondrial humano: considerado en su conjunto (16,5 kb), sus dos hebras son desequilibradas, pues una es más rica en purinas que la otra (56 y 44%, respectivamente). A la primera se la llama habitualmente "hebra pesada" (o hebra "H", del inglés *heavy*; recuérdese el mayor tamaño del anillo purínico) y a la otra, "hebra ligera" (o "L", de *light*).

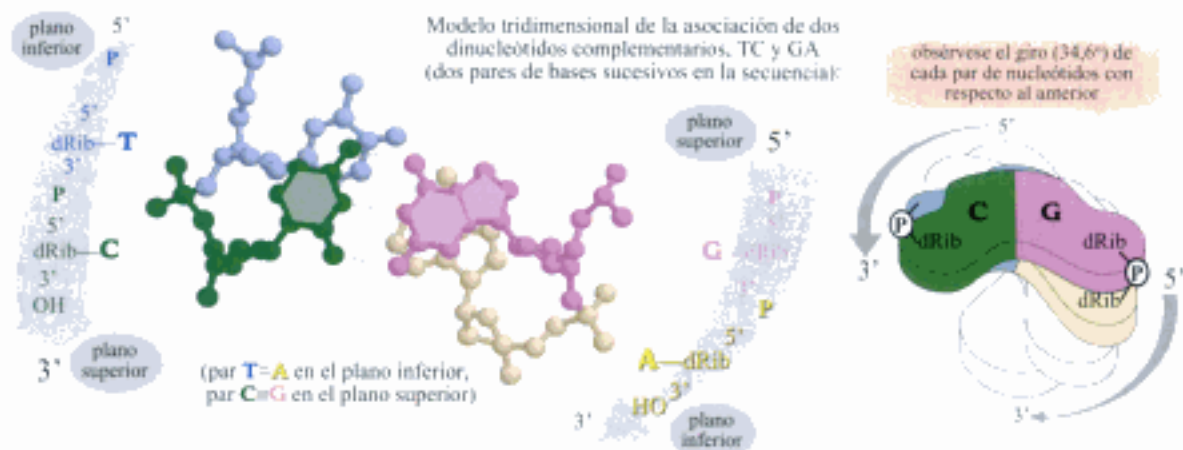
5.2.5 Helicidad y carácter dextrógiro

Las restricciones estéricas impuestas por la geometría de los pares de bases (que incluye los enlaces entre base, pentosa y fosfato, pág. 43) hacen que el apareamiento completo de las dos hebras sólo se pueda establecer si los pares de bases sucesivos van girando unos con respecto a otros (concretamente, $34,6^\circ$ en el B-DNA). Como consecuencia de la acumulación de estos giros, las hebras de DNA adoptan una disposición helicoidal (es decir, cada hebra describe una **hélice**) y la cadena de doble hebra es una "**doble hélice**".

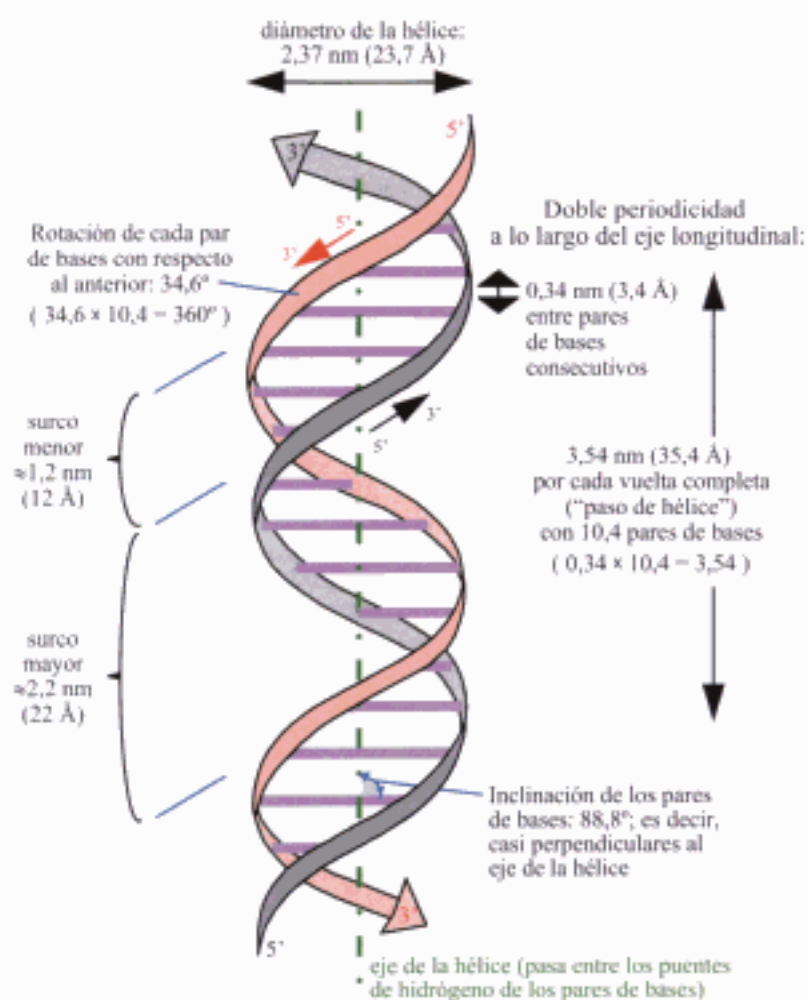
Geométricamente, una hélice (genérica, no de DNA), sea sencilla o doble, puede ser de dos tipos:



En el caso de la forma B del DNA ambas hebras se enrollan alrededor de un eje común, longitudinal, formando una **doble hélice dextrógiro**. La observación de modelos moleculares tridimensionales permite una mejor comprensión de este carácter helicoidal.



5.2.6 Parámetros cuantitativos de la doble hélice



Ejercicio:

A partir del dato de separación entre cada dos pares de bases (0,34 nm), se puede calcular el número total de pares de bases de una molécula de B-DNA de cualquier longitud (por ejemplo, un metro):

$$1 \text{ m} / 0,34 \text{ nm/pb} = 10^9 \text{ nm} / 0,34 \text{ nm/pb} = 2,94 \cdot 10^9 \text{ pb} \approx 3.000 \text{ millones de pb (Mb)}$$

Viceversa, a partir del número de pares de bases de un DNA (p.ej., 5.880 Mb), frecuentemente llamado su "longitud en pares de bases", se puede calcular la longitud (dimensiones) de la molécula:

$$5.880 \text{ Mb} = 5.880 \cdot 10^6 \text{ pb} = 5,88 \cdot 10^9 \text{ pb}$$

$$5,88 \cdot 10^9 \text{ pb} \cdot 0,34 \text{ nm/pb} = 2 \cdot 10^9 \text{ nm} = 2 \cdot 10^6 \text{ } \mu\text{m} = 2 \cdot 10^3 \text{ mm} = 2 \text{ m}$$

(nota: $10^6 \text{ pb} = 1$ millón de pb = 1 Mpb, pero se suele usar Mb y no Mpb)

La escalera Francisco I del castillo de Chambord (valle del Loira, Francia), cuyo diseño se atribuye a Leonardo da Vinci, ha sido considerada como modelo de la doble hélice del DNA, con la salvedad de que su giro es levógiro.

Se trata de una escalera helicoidal (de tipo caracol), con la característica atípica de poseer dos rampas independientes, cada una con su entrada (en la foto, una al frente y otra detrás). Por tanto, dos personas pueden utilizarla al mismo tiempo sin encontrarse.

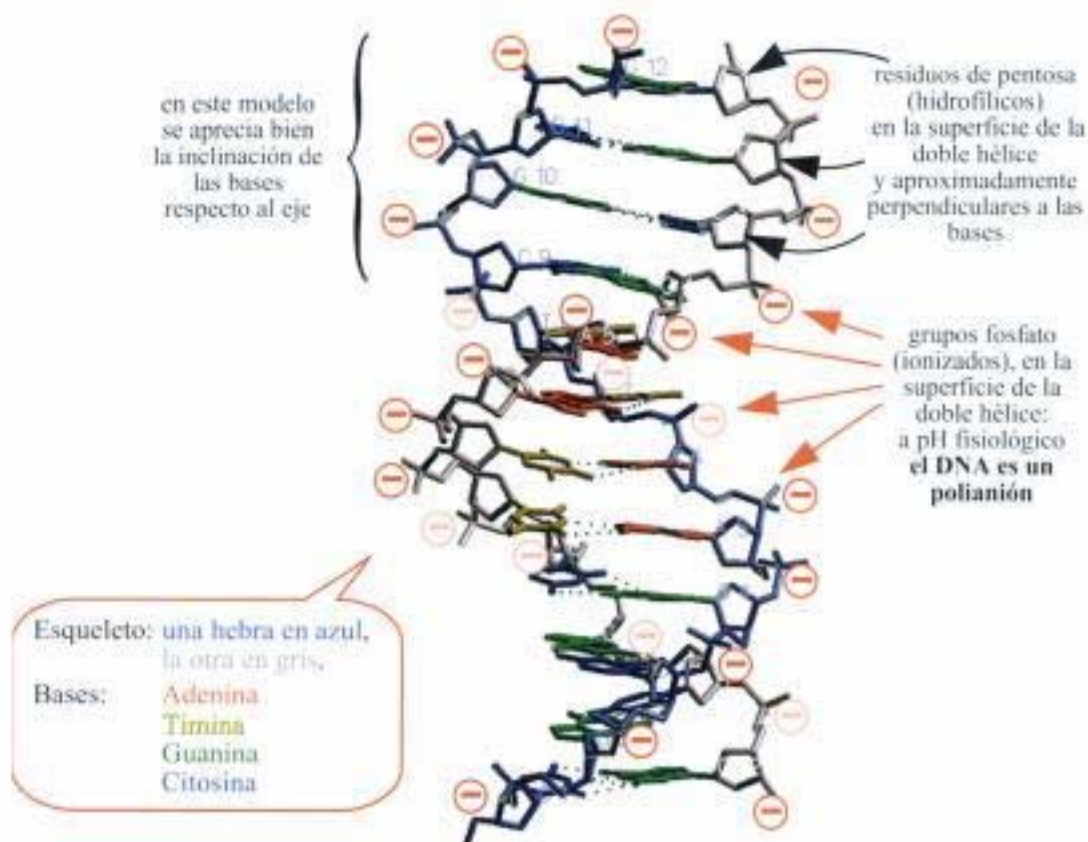
En el caso del B-DNA, el esqueleto de las dos hebras correspondería a las barandillas externas de las dos rampas.



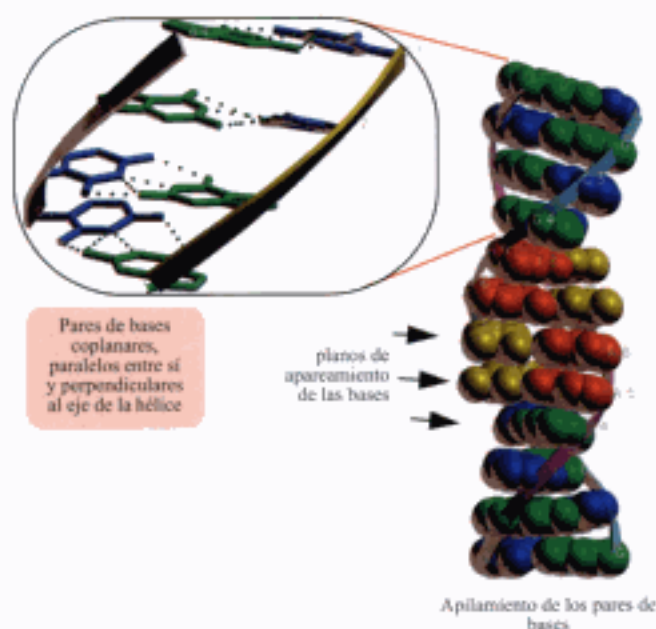
5.2.7 Carácter anfipático y estabilidad de la doble hélice

5.2.7.1 Situación de los fosfatos y carácter hidrofílico

Debido a la interacción entre las bases de las dos hebras, los esqueletos azúcar-fosfato de ambas deben disponerse hacia el exterior de la doble hélice, por lo que el DNA dúplex muestra un marcado carácter hidrofílico. La superficie de la molécula establece así interacciones favorables con el medio acuoso circundante. Al mismo tiempo, la distribución relativa de los fosfatos es tal que minimiza la repulsión electrostática entre ellos.

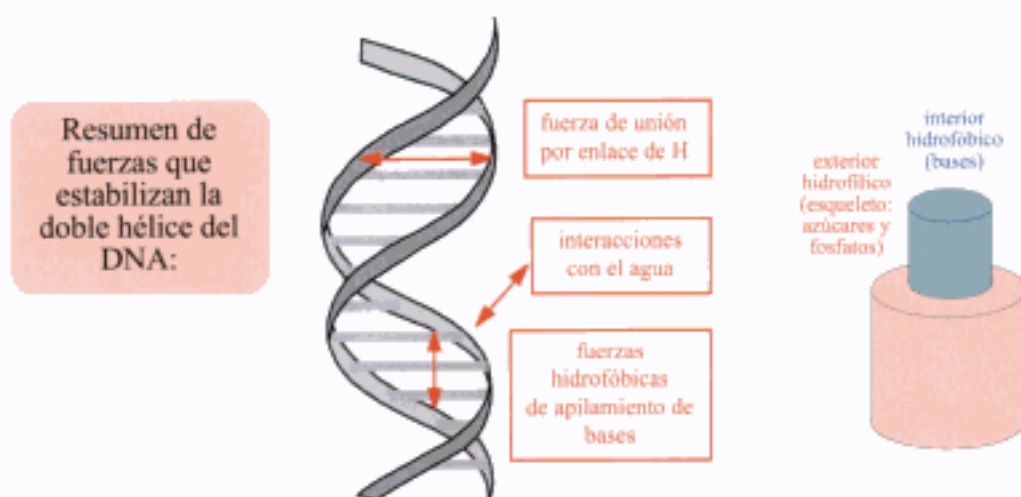


5.2.7.2 Coplanariedad, apilamiento de las bases y carácter hidrofóbico



Las bases purínicas y pirimidínicas de ambas cadenas se encuentran dirigidas hacia el interior de la doble hélice, superpuestas, formando un apilamiento similar a un montón de monedas. Como consecuencia de la proximidad y paralelismo de los anillos aromáticos, se originan entre ellos fuerzas de atracción, de tipo Van der Waals, llamadas *interacciones de apilamiento* (fuerzas hidrofóbicas). De esta forma, el interior de la doble hélice, altamente hidrofóbico, queda aislado del medio acuoso (hidrofílico), contribuyendo a la estabilización de la molécula y a la protección de la información genética (pues ésta la constituye la secuencia de bases).

Estas interacciones son mayoritariamente inespecíficas (es decir, son independientes del tipo de base implicada), a diferencia de los enlaces de hidrógeno entre bases complementarias, muy específicos. Aunque la atracción entre dos pares de bases individualmente no es muy fuerte, al existir 10,4 bases por vuelta de hélice, la contribución total a la estabilidad de la molécula es muy elevada.

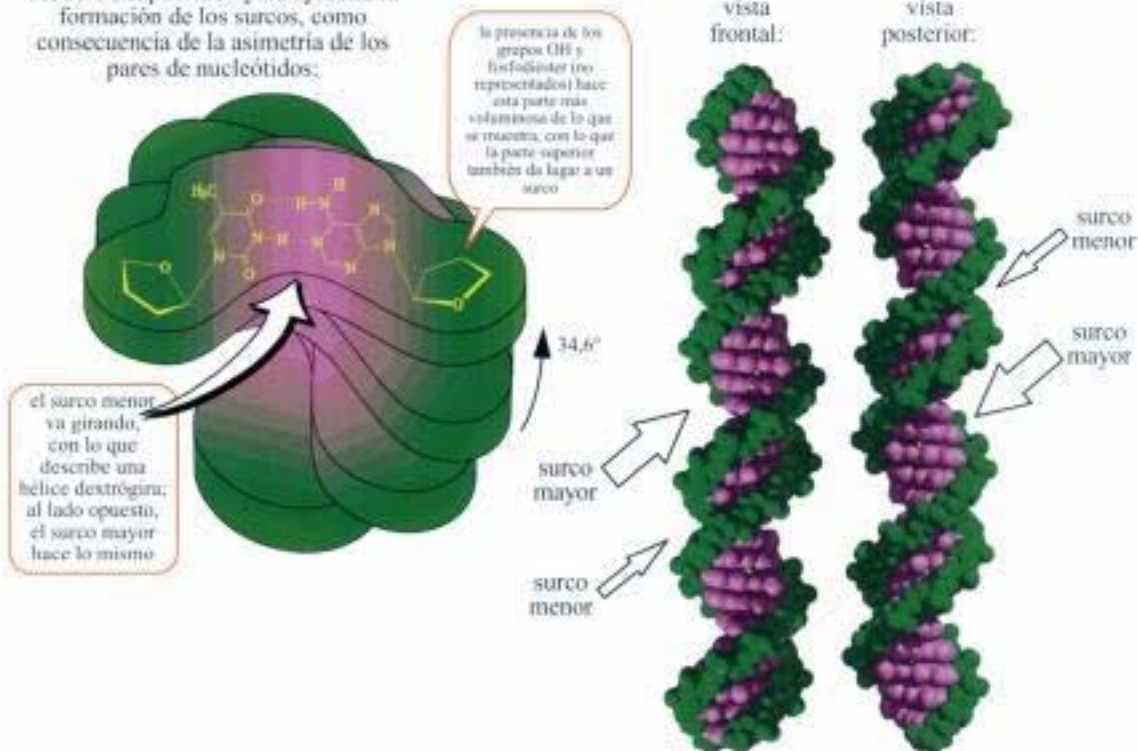


5.2.8 Superficie de la molécula: surcos en el DNA

Como se ha observado, la geometría de los pares de bases (pág. 42) y el antiparalelismo de las hebras (pág. 43), necesarios para la formación de los puentes de hidrógeno, hacen que los C-1' de los azúcares de los dos nucleótidos complementarios no se dispongan simétricamente con respecto a las bases y, por ello, que los esqueletos desoxirribosa-fosfato de las dos cadenas no se sitúen en puntos diametralmente opuestos de la doble hélice. Como consecuencia, a todo lo largo de la estructura en doble hélice se forman dos surcos, uno mayor (o hendidura profunda) y otro menor (o hendidura pequeña), que van girando de forma helicoidal, al igual que los dos esqueletos. Los surcos, especialmente el mayor, dejan espacio suficiente para que moléculas externas puedan contactar con las bases; éste es el fundamento de la interacción del DNA con numerosas proteínas capaces de reconocer secuencias de bases (pág. 60), con funciones tan importantes como la regulación de la expresión génica (por ejemplo, factores de transcripción). En cierto modo, es como si las proteínas "leyeran" la secuencia a través de los grupos funcionales de las bases que "asoman" dentro del surco mayor del DNA.

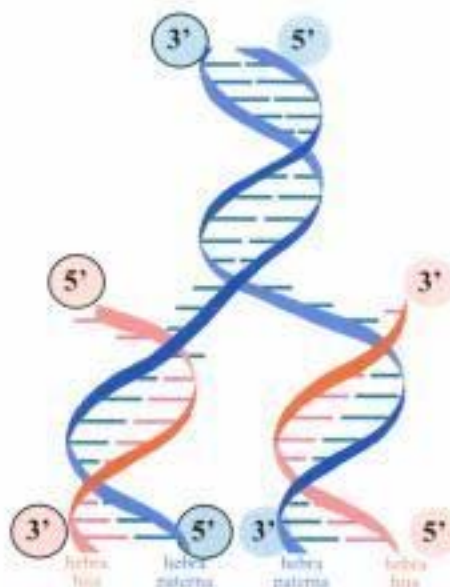


Modelo simplificado para apreciar la formación de los surcos, como consecuencia de la asimetría de los pares de nucleótidos:



5.2.9 Significado biológico: idoneidad para la replicación

El modelo de doble hélice de Watson y Crick permite explicar que se mantenga una estructura regular a pesar de la variación de secuencia y que se produzca, mediante un mecanismo muy simple, la duplicación del material genético. En resumen, la replicación procede de una manera lógica, con dos procesos: (a) la separación de las dos hebras, y (b) la síntesis, a partir de ellas como molde, de otras dos cadenas complementarias (los detalles de este proceso se estudiarán posteriormente, tema 12). También permite la reparación de mutaciones y daños en una hebra, al disponerse de la otra como referencia de la secuencia correcta.



Tema 6

Variaciones en la estructura del DNA

6.1	Variantes en doble hebra: formas A y Z	52
6.1.1	Forma A del DNA, en comparación con la forma B	52
6.1.2	Forma Z del DNA, en comparación con las formas B y A	55
6.2	Variantes locales de la estructura secundaria del B-DNA	56
6.2.1	Curvatura de la doble hélice	57
6.2.2	Palíndromos	57
6.2.2.1	Concepto y características	57
6.2.2.2	Repeticiones de secuencia	58
6.2.2.3	Palíndromos en moléculas de hebra sencilla	59
6.2.2.4	Consecuencias estructurales de los palíndromos	59
	a) En ácidos nucleicos de hebra sencilla	59
	b) En DNA de doble hebra	59
6.2.2.5	Funcionalidad de los palíndromos	59
6.2.3	H-DNA: triple hélice	60
6.3	Motivos estructurales responsables de la unión del DNA con proteínas	60
6.3.1	Descripción de los motivos estructurales más frecuentes	61
6.3.1.1	Motivo hélice-giro-hélice o hélice-vuelta-hélice	61
6.3.1.2	Motivo hélice-bucle-hélice	62
6.3.1.3	Motivo homeodominio	62
6.3.1.4	Motivo dedo de zinc	63
6.3.1.5	Motivo cremallera de leucina	64

La molécula de DNA no es una estructura estática, sino flexible y dinámica. Gracias a la capacidad de rotación alrededor de los enlaces de los nucleótidos, el DNA puede adoptar *in vivo* **diversas formas**, distintas de la forma B de Watson y Crick. Se estudian en este tema:

- Aquéllas que, manteniendo la estructura helicoidal de doble hebra, presentan una conformación diferente: formas A y Z.
- Las variaciones conformacionales de carácter local en torno a la estructura B.
- La interacción de regiones específicas del DNA con estructuras características de algunas proteínas.

En general, estos tres tipos de variaciones estructurales tienen relación con funciones importantes en el metabolismo del DNA, interviniendo, por ejemplo, en el inicio y la regulación de los procesos de replicación, recombinación, transcripción, etc. De ahí el interés de su estudio conjunto, como tema independiente, para ampliar la estructura secundaria y sentar las bases de dichas funciones, que serán objeto de estudio en temas posteriores.

6.1 VARIANTES EN DOBLE HEBRA: FORMAS A Y Z

El análisis por difracción de rayos X a resolución atómica reveló la existencia de dos nuevas formas estructurales o conformaciones, diferentes a la B-DNA tradicional, en las que se mantienen las características básicas de éste (complementariedad, antiparalelismo, helicidad, etc.), pero presentan otras diferencias:

Comparación entre los tres tipos estructurales básicos del DNA en doble hebra:

	TIPO DE HÉLICE		
	A (deshidratada)	B (Watson-Crick)	Z (Rich-Dickerson)
Sentido de giro de la hélice	Dextrógiro	Dextrógiro	Levógiro
Forma y tamaño	La más ancha y corta	Intermedia	La más estrecha y larga
Diámetro de la hélice	2,55 nm (25,5 Å)	2,37 nm (23,7 Å)	1,84 nm (18,4 Å)
Distancia entre bases (d) ¹	0,23 nm (2,3 Å)	0,34 nm (3,4 Å)	0,38 nm (3,8 Å)
Pares de bases por vuelta (n)	11	10,4	12
Rotación por cada base ($360^\circ/n$) ²	+ 32,7°	+ 34,6°	- 30°
Longitud vuelta de hélice (n/d) ³	2,53 nm (25,3 Å)	3,54 nm (35,4 Å)	4,56 nm (45,6 Å)
Inclinación del plano de los pares de bases ⁴	19° (gran inclinación)	1,2° (casi perpendicular al eje de la hélice)	9° (ligera inclinación)
Surco mayor o grande	Estrecho, profundo	Ancho, profundidad media	Plano, sin profundidad
Surco menor o pequeño	Amplio, no profundo	Estrecho, profundidad media	Estrecho, profundo
Enlace N-glicosídico	Anti	Anti	Sin (C,T) y Anti (G)

Notas: ¹ Distancia entre pares de bases, medida a lo largo del eje de la hélice.

² Rotación por cada par de bases, o giro respecto al par anterior (+ dextrógiro, - levógiro).

³ Longitud de la vuelta de hélice, medida a lo largo del eje; coloquialmente "paso de rosca" o "paso de hélice".

⁴ Inclinación respecto al plano perpendicular al eje de la hélice.

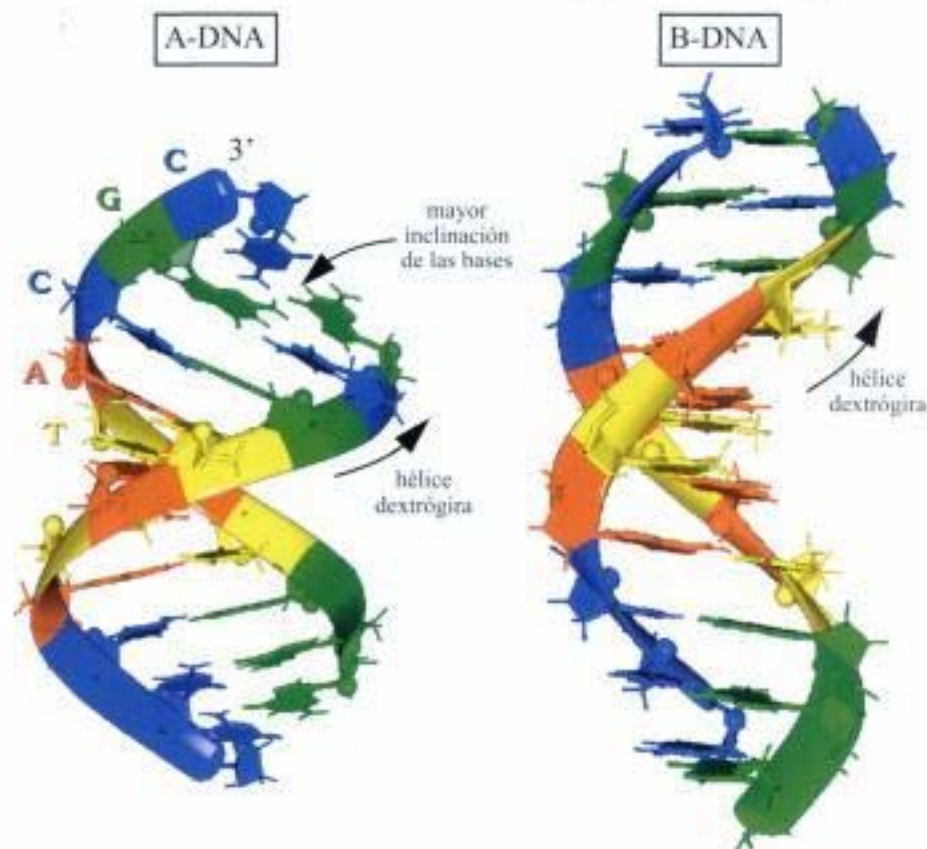
6.1.1 Forma A del DNA, en comparación con la forma B

La reducción de la humedad relativa por debajo del 75% (por ejemplo, en disoluciones muy concentradas, es decir, deshidratadas) hace que el B-DNA tienda a transformarse, como ejemplo de la flexibilidad de la molécula, a otra conformación, denominada A-DNA. Aunque esta forma no se ha encontrado bajo condiciones fisiológicas, su estudio es de interés por ser la que adopta el RNA cuando posee regiones de doble hebra, así como la de los híbridos DNA-RNA.

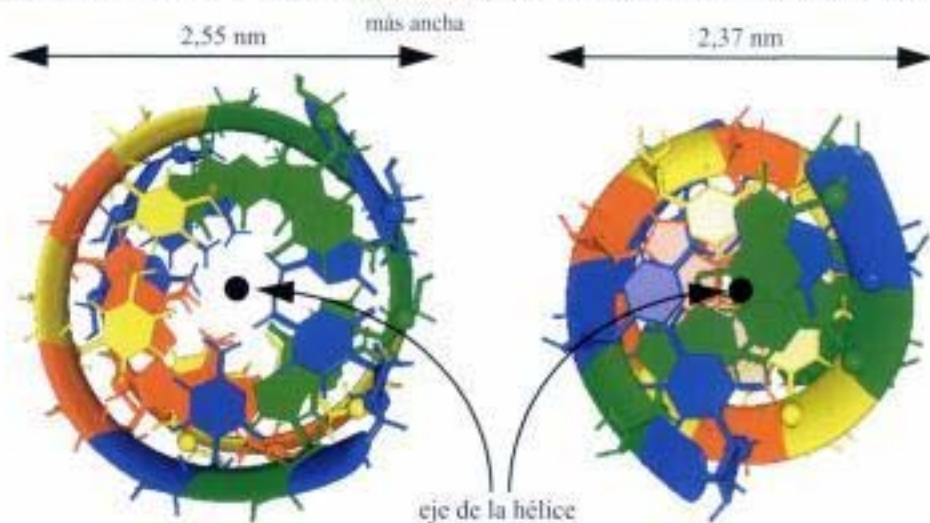
Estructuralmente, el A-DNA es similar a la forma B: es igualmente una doble hélice dextrógira, con hebras complementarias y antiparalelas unidas por enlaces de hidrógeno. Sin embargo, difiere en sus dimensiones: la hélice A es **más ancha** (mayor diámetro), **más corta** para un mismo número de nucleótidos (menor distancia entre pares de bases) y tiene **mayor número de bases por vuelta**. Como consecuencia, son menores la rotación por cada base y el paso de hélice. Además, las bases no se disponen casi perpendicularmente al eje de la hélice, sino claramente inclinadas, y se sitúan más separadas del eje de la hélice. La superficie de la molécula es también diferente: el surco mayor es más estrecho que en la forma B y muy profundo, mientras que el surco menor casi no se aprecia por ser ancho y poco profundo. Mientras que la forma B puede acomodar en su surco menor una columna de moléculas de agua, el surco superficial de la forma A no ofrece espacio suficiente; esta diferencia explica, al menos en parte, por qué la deshidratación favorece la forma A.

Para mostrar las diferencias estructurales se puede emplear un **modelo de "cintas"**, donde el esqueleto se representa por una cinta, de la cual se destacan los **anillos pentagonales** de desoxirribosa, sus **átomos de oxígeno** (esferas) y los **grupos fosfato** (varillas que sobresalen). Cada nucleótido está representado en un color diferente.

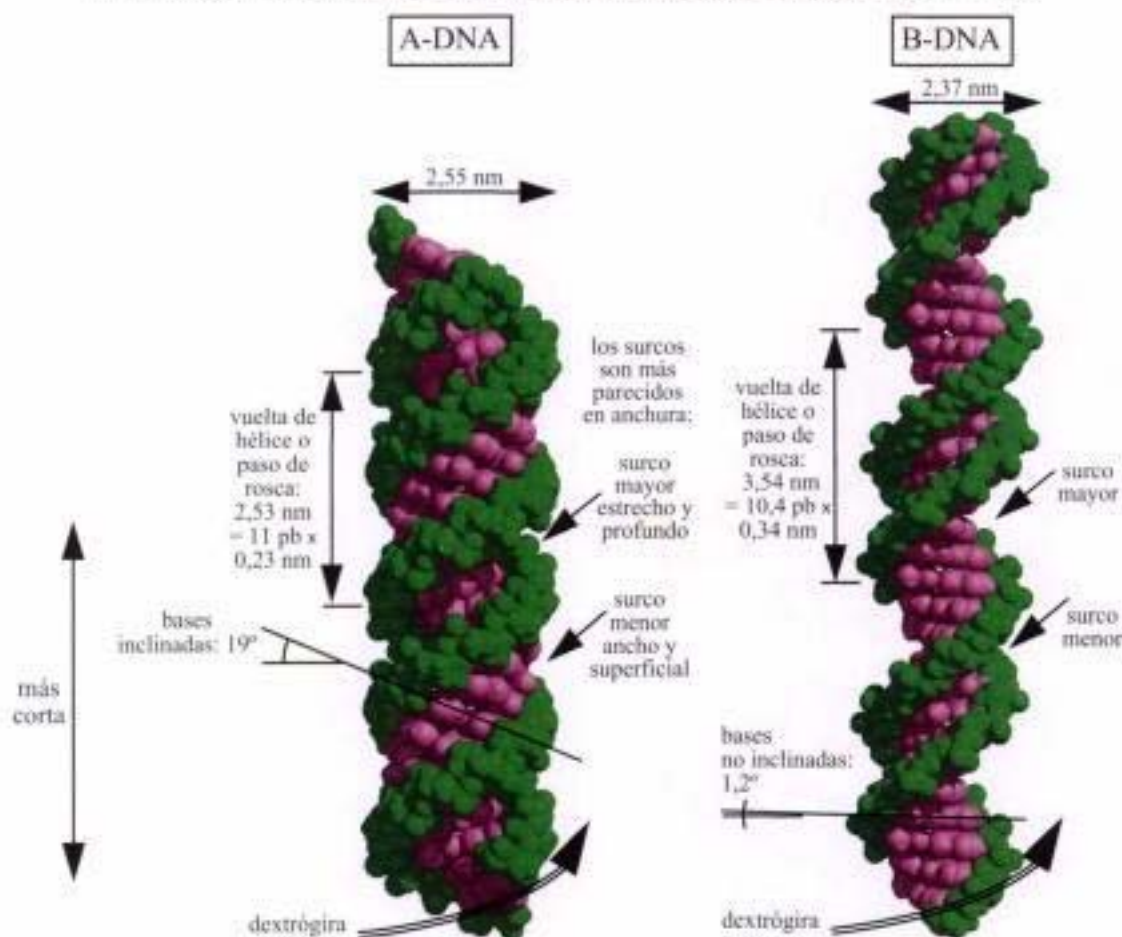
La secuencia en ambos casos es la misma: (5') GGGTATACGC (3')
(3') CCCATATGCG (5')



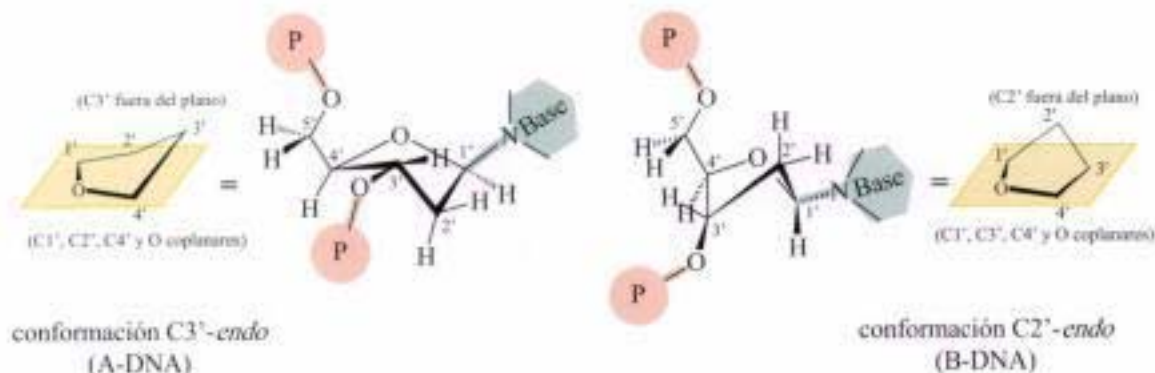
Este mismo modelo (visto desde su parte superior) permite ilustrar que la hélice A es más abierta, con mayor diámetro y una disposición de las bases más alejada del eje de la hélice:



Una visión complementaria a la anterior, a la vez que más cercana a la realidad, es la del **modelo espacial compacto**. Las bases están coloreadas en rosa y el esqueleto de pentosas y fosfatos en verde.



Las diferencias, recién comentadas, en la estructura de las dobles hélices A y B vienen originadas en última instancia por la diferente disposición de los pares de bases, debida concretamente a una **distinta conformación de la desoxirribosa**: En el A-DNA, el C3' sobresale del plano formado por los otros cuatro carbonos de la ribofuranosa (conformación C3'-endo), mientras que en el B-DNA es el C2' el que no es coplanar (conformación C2'-endo).



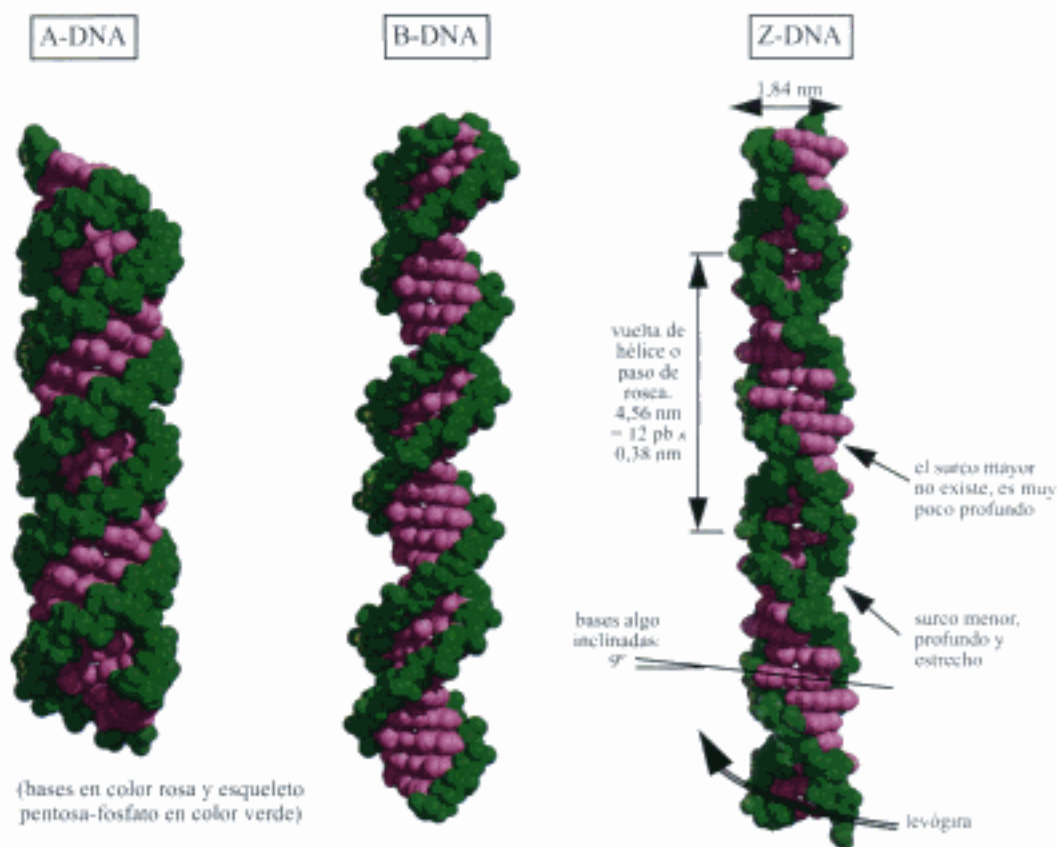
Esta distinta conformación del anillo de ribosa hace que en el A-DNA los dos grupos fosfato consecutivos estén más próximos entre sí, lo que explica que los pares de bases sucesivos se acerquen, acortando la cadena. Similarmente, el cambio en la geometría conduce a los diferentes parámetros geométricos ya comentados.

6.1.2 Forma Z del DNA, en comparación con las formas B y A

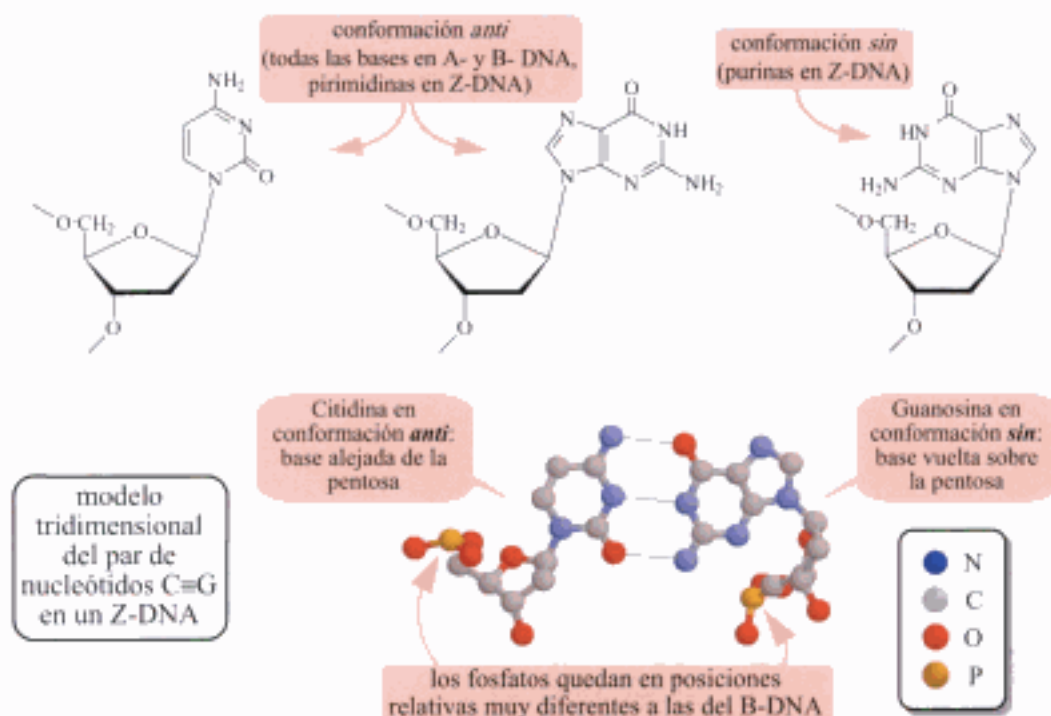
La existencia de la forma Z surge por primera vez de la observación, mediante difracción de rayos X, de una doble hélice del hexanucleótido CGCGCG, con **características distintas** a las formas B y A: entre ellas, como muy significativa, su giro a izquierdas o levógiro. Posteriormente, se ha encontrado esta estructura en otros oligonucleótidos sintéticos con secuencias de purinas y pirimidinas alternas. Aunque *in vivo* predomina el B-DNA, se ha podido observar la presencia de algunas regiones Z en el genoma de células eucarióticas, empleando anticuerpos contra esta última conformación. Su función, aún por confirmar, parece estar relacionada con el control de la expresión génica y con la recombinación.

En cuanto a sus dimensiones, observadas ahora comparativamente con las otras conformaciones, se observa que la hélice Z es la **más estrecha** (1,84 nm de diámetro), la **más larga** (mayor distancia entre bases: 0,38 nm) y la que tiene **mayor número de pares de bases por vuelta** (12), lo cual supone una menor rotación por base (30°) y un mayor paso o longitud de una vuelta de hélice (4,56 nm). Es, en resumen, una hélice muy estirada. En su superficie forma **un solo surco, profundo** (correspondiente a la posición del surco menor en A- y B-DNA).

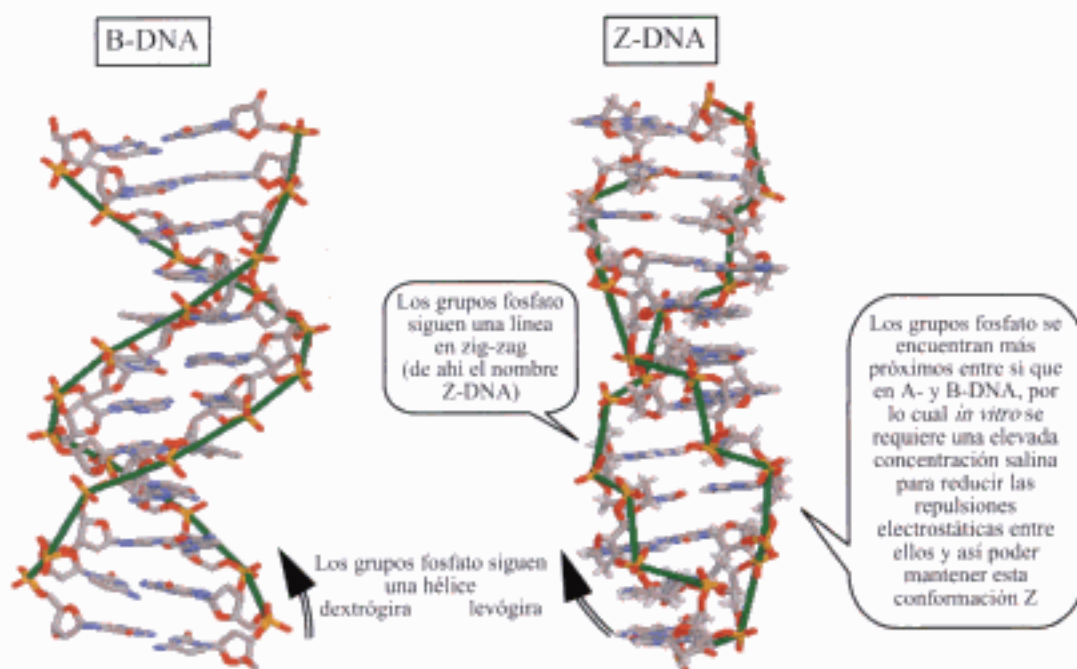
Modelos espaciales compactos:



La consideración de esta estructura a nivel molecular muestra que sus diferencias respecto a las formas A y B se deben también a una **distinta conformación**, en este caso alrededor del enlace N-glicosídico. En las formas A y B todas las bases nitrogenadas se orientan alejadas del anillo de pentosa, en la llamada conformación **anti**, mientras que en el Z-DNA sólo lo hacen las pirimidinas; las purinas tienden a situarse sobre la pentosa (conformación **sin**). Por ello, en las secuencias de purinas y pirimidinas alternantes (por ejemplo, CGCGCGCG), típicas del Z-DNA, se favorece la sucesión de ambas conformaciones, *sin* (G) y *anti* (C), obligando al esqueleto azúcar-fosfato a plegarse en zig-zag para poder establecer los enlaces de hidrógeno.



La disposición en zig-zag del esqueleto de la forma Z del DNA se aprecia de forma óptima en este modelo de varillas, donde los grupos fosfato se han unido con una línea verde hipotética. Se muestra, a efectos comparativos, cómo el B-DNA no presenta esta ordenación, sino que su esqueleto describe una hélice continua. Obsérvese que el Z-DNA es el único con hélice en sentido levógira.



6.2 VARIANTES LOCALES DE LA ESTRUCTURA SECUNDARIA DEL B-DNA

Los estudios de difracción de rayos X también revelaron la existencia de variaciones en la conformación en determinadas secuencias del B-DNA. Aunque dentro de este apartado también podría incluirse el superenrollamiento del DNA, se ha preferido describir este aspecto como una estructura de orden superior de los ácidos nucleicos, dada su implicación en la organización del cromosoma.

6.2.1 Curvatura de la doble hélice

A pesar de la regularidad del modelo de Watson y Crick, existen pequeñas diferencias locales en la estructura de los DNAs dependiendo de los nucleótidos integrantes. Por ejemplo, en determinadas secuencias los pares de bases no son exactamente coplanares, lo que puede ocasionar una desviación del eje de la doble hélice. En otros casos, frecuentes, la curvatura no se origina por una secuencia particular de nucleótidos, sino que es forzada por la unión de una proteína a la cadena de DNA (como, por ejemplo, la mostrada en pág. 61)

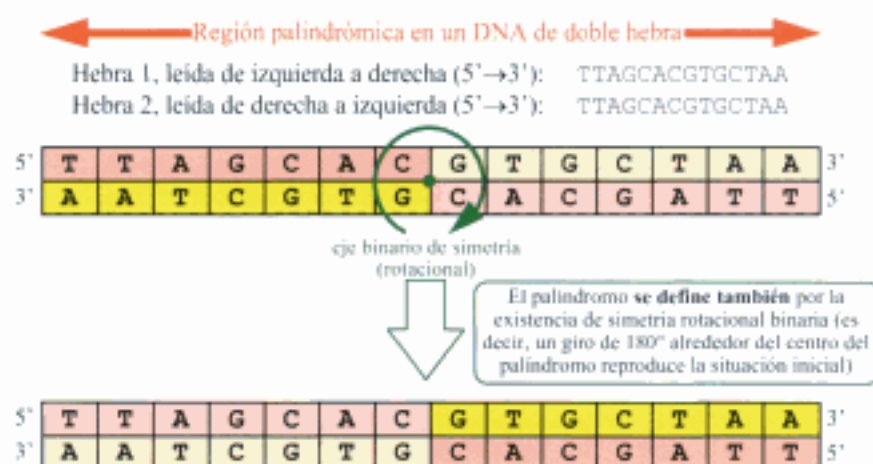
6.2.2 Palíndromos

6.2.2.1 Concepto y características

Se denomina secuencias palindrómicas, o simplemente palíndromos, a aquellas regiones de un DNA cuya secuencia es la misma en ambas hebras. Son bastante comunes, y tienen un significado especial por varias razones: (a) ser lugares-diana de las enzimas de restricción (herramientas experimentales para escindir ácidos nucleicos en la tecnología del DNA recombinante); (b) actuar como centros reguladores de la expresión génica; (c) dar lugar a estructuras secundarias peculiares en DNA y RNA, etc.

Gramaticalmente, un **palíndromo** (del griego *palin* = de nuevo, *dromos* = carrera) es toda palabra o frase que se deletrea igual de izquierda a derecha y de derecha a izquierda. Por ejemplo: RADAR, ANILINA, DABALE ARROZ A LA ZORRA EL ABAD. En los palíndromos hay un centro de simetría (a la vez eje y plano de simetría), que es la letra central o la posición entre las dos letras centrales.

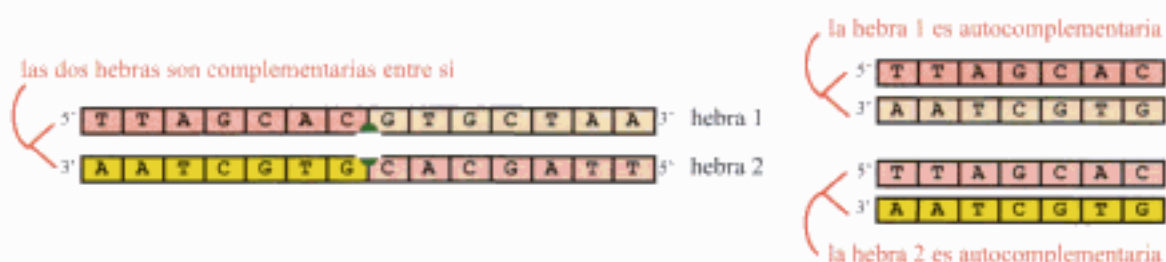
Sin embargo, la extensión de este término al campo de la Biología Molecular no es directa: se dice que una región de DNA es palindrómica cuando la secuencia de **una hebra**, leída de izquierda a derecha, es igual que la de **la otra hebra**, leída de derecha a izquierda. Por tanto, los palíndromos se observan en las dos cadenas conjuntamente, y no en una sola. Por ejemplo, la secuencia TTAGCAC·CACGATT sería un palíndromo gramatical, pero no es una secuencia palindrómica. Teniendo en cuenta que, según el convenio de escritura de los DNAs, ambas hebras se deben leer de 5' a 3', se concluye que **las dos hebras de un palíndromo tienen la misma secuencia**; ésta puede considerarse una definición de palíndromo.



Como una extensión del concepto de palíndromo, se llaman **palíndromos interrumpidos** aquéllos en los cuales una pequeña secuencia separa las dos mitades del palíndromo y actúa, por tanto, como centro de simetría no puntual (en lugar del espacio entre dos nucleótidos de un palíndromo estricto). En cualquier caso, todo palíndromo está formado por dos "mitades" de secuencia idéntica, separadas por un centro de simetría, puntual o no puntual.



Al cumplirse tanto la simetría propia del palíndromo como la complementariedad de las dos hebras resulta que las secuencias de una misma hebra que quedan a ambos lados del centro de simetría son complementarias entre sí; se dice por ello que las secuencias palindrómicas son **autocomplementarias**.



6.2.2.2 Repeticiones de secuencia

En ocasiones se utiliza como sinónimo de palíndromo el término *repeticiones invertidas*, puesto que si se leen en sentidos opuestos, las dos mitades tienen la misma secuencia (a ambos lados del centro de simetría, puntual o no).

Otro término relacionado es el de *repeticiones directas*, que se emplea cuando en un DNA hay repetición de una secuencia (unidad de repetición), leída siempre en el mismo sentido. No son palíndromos porque no se cumple el criterio de igualdad de secuencias leídas en sentido opuesto, ni hay centro de simetría, por lo que no se habla de "mitades". Estas repeticiones tienen interés porque aparecen con enorme frecuencia en el genoma (pág. 109), tanto en posiciones contiguas o adyacentes (*repeticiones en tándem*) como separadas por un número elevado de pares de bases (*repeticiones dispersas*).

Finalmente, se habla de *repeticiones especulares* cuando la secuencia se repite, de forma invertida, dentro de la misma hebra. De nuevo, no son palíndromos, pues no hay centro de simetría rotacional, sino plano de simetría especular.

Estas dos últimas posibilidades, aunque aparecen con frecuencia en los DNAs, ocupando una parte importante de su secuencia, no tienen consecuencias en cuanto a la estructura secundaria de los ácidos nucleicos, por lo que a partir de este momento sólo se tratarán los palíndromos.

palíndromos = repeticiones invertidas:

(ejemplo: palíndromo discontinuo, con un centro de simetría de 4 pb)



Las secuencias son iguales si se leen en sentidos opuestos:

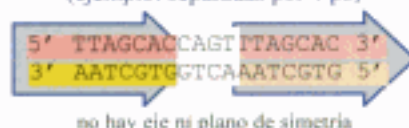
Secuencia 1ª mitad (→): Secuencia 2ª mitad (←):

5' TTAGCAC 3' = 5' TTAGCAC 3'

AATCGTG AATCGTG

repeticiones directas:

(ejemplo: separadas por 4 pb)



Las secuencias son iguales si se leen en el mismo sentido:

Secuencia 1ª repetición (→): Secuencia 2ª repetición (→):

5' TTAGCAC 3' = 5' TTAGCAC 3'

AATCGTG AATCGTG

repeticiones especulares:

(ejemplo: separadas por 4 pb)



Las secuencias son diferentes:

Secuencia 1ª repetición (→): Secuencia 2ª repetición (←):

5' TTAGCAC 3' ≠ 5' AATCGTG 3'

AATCGTG TTAGCAC

Sólo iguales si la segunda se lee en sentido no convencional, de 3' a 5' :

5' TTAGCAC 3' = 3' TTAGCAC 5'

AATCGTG AATCGTG

6.2.2.3 Palíndromos en moléculas de hebra sencilla

Para el caso de ácidos nucleicos de hebra sencilla, la definición de palíndromo no se puede aplicar estrictamente. Sin embargo, dada la relación de complementariedad entre la secuencia de un RNA y la de una de las hebras del DNA que lo codifica, las secuencias de RNA de hebra sencilla codificado por un DNA palindrómico son también autocomplementarias y se consideran como palíndromos. Bajo este criterio, cada una de las dos hebras de un DNA palindrómico también es una secuencia palindrómica de hebra sencilla.

6.2.2.4 Consecuencias estructurales de los palíndromos

a) En ácidos nucleicos de hebra sencilla

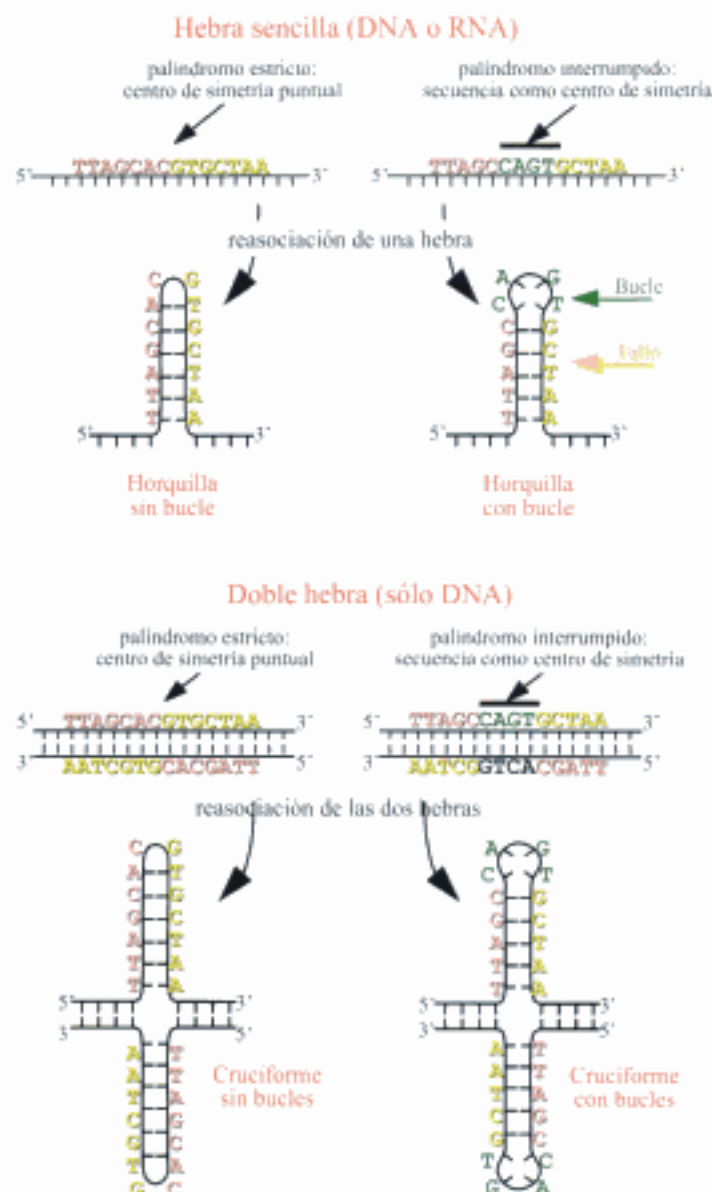
La existencia de palíndromos, por ser secuencias autocomplementarias, puede afectar de forma importante a la estructura secundaria. Así, las hebras sencillas de DNA o RNA pueden plegarse ocasionalmente sobre sí mismas para formar una estructura en horquilla. Si el centro de simetría no es puntual (palíndromos interrumpidos), su secuencia forma un *bucle* en el extremo de la horquilla.

La formación de horquillas posee gran relevancia en la estructura tridimensional de algunos RNAs. El ejemplo más significativo lo constituyen los brazos de los tRNAs, regiones con estructura de *tallo y bucle* (pág. 74), formados por secuencias palindrómicas. Por otro lado, este tipo de plegamiento en el RNA contribuye, en algunos casos, a la terminación de la transcripción; concretamente, en el mecanismo de terminación independiente de ρ de procariotas y los mecanismos equivalentes de eucariotas.

En el caso de hebras sencillas de DNA también pueden encontrarse horquillas, por ejemplo en el DNA desnaturalizado, que carece de una estructura secundaria definida. También se propone su formación en situaciones particulares, por ejemplo durante la síntesis en el laboratorio del cDNA (DNA complementario) a partir de un mRNA, donde el extremo 3' del DNA de hebra sencilla se repliega sobre sí mismo y sirve así de cebador para la polimerasa (pág. 206).

b) En DNA de doble hebra

En el caso de un DNA doble, las hebras del palíndromo pueden separarse y reasociarse de otra forma, generando estructuras de doble horquilla o *cruciformes* (en forma de cruz). Similarmente al caso anterior, los palíndromos interrumpidos forman cruciformes con dos *bucles*.



6.2.2.5 Funcionalidad de los palíndromos

Los palíndromos en DNA dúplex actúan como señales de reconocimiento a las que se unen numerosas proteínas reguladoras, que poseen estructuras dimericas que les permiten unirse de forma simétrica a ambas mitades del palíndromo (pág. 61).

6.2.3 H-DNA: triple hélice

Se trata de una estructura poco habitual, formada por 3 hebras: una **triple hélice**, también llamada *triplex* o *tríplice*. Se descubrió por primera vez en un RNA, aunque es en el DNA donde se encuentra con más frecuencia. Se forman en regiones ricas en pirimidinas (secuencia $(CT)_n$) en una hebra y ricas en purinas (secuencia $(AG)_n$) en la hebra complementaria. En condiciones normales, dichas secuencias se encuentran totalmente apareadas sin alteración alguna en la doble hélice. Sin embargo, puede ocurrir que parte de la doble hélice se abra y la hebra rica en C y T se repliegue y se aparee con la otra hebra (AG), mediante una nueva clase de enlaces de hidrógeno (conocidos como enlaces de Hoogsteen), en una zona donde ésta aún forma doble hélice. Aparece así una triple hélice $(CT)_n / (AG)_n / (CT)_n$ en dicho segmento del DNA. La otra hebra sencilla (rica en A y G), desapareada, forma una especie de bucle.

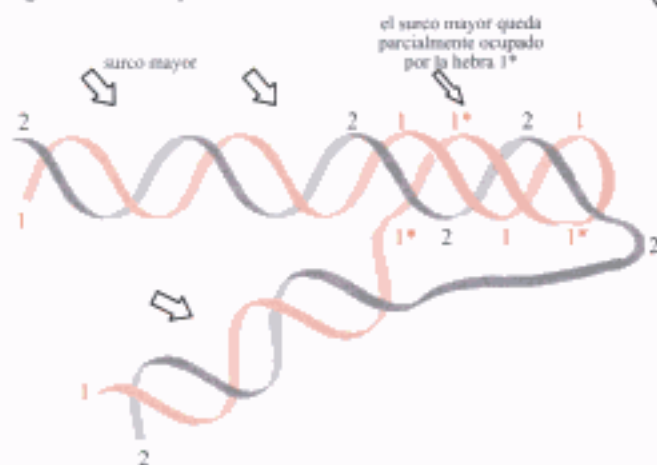
Organización de las cadenas en el plano:



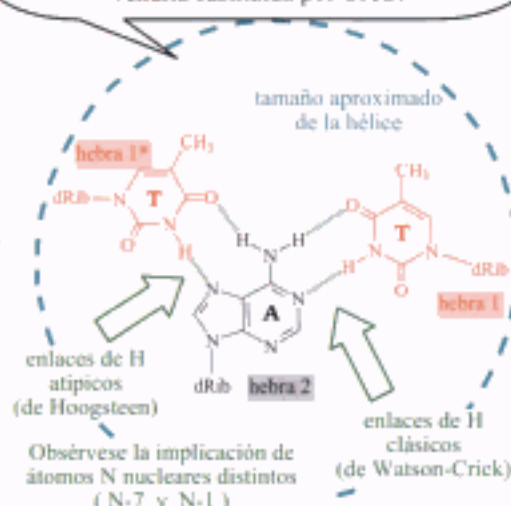
Esquemáticamente:

- 1º) Las dos hebras (1 y 2) forman una doble hélice.
- 2º) Una de las hebras (2) gira volviendo hacia atrás (1*) y se aparea con las dos hebras anteriores, formando el segmento de triple hélice.
- 3º) La otra hebra (2) queda desapareada, hasta que se reúne de nuevo con la hebra 1 para formar un segundo tramo de doble hélice.

Organización espacial de las cadenas:



Ejemplo de apareamiento de 3 bases que da lugar a la triple hélice en el H-DNA (TAT). El otro apareamiento es entre citosina, guanina y citosina protonada: CGC⁺. En el caso de RNAs, la interacción TAT vendría sustituida por UAU.



Este tipo de estructura ha despertado gran interés por su aplicación en terapias de inhibición de la expresión génica, mediante la unión de un oligonucleótido sintético a la doble hélice del DNA en la región promotora de un gen, formando una triple hélice que impide la expresión del gen.

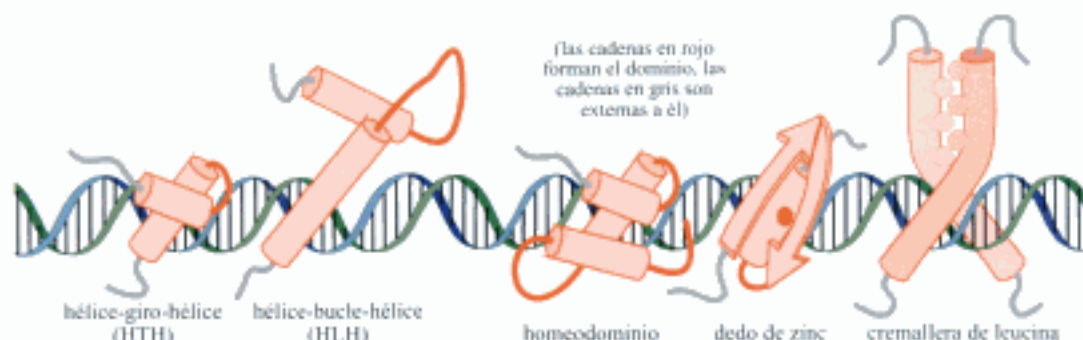
6.3 MOTIVOS ESTRUCTURALES RESPONSABLES DE LA UNIÓN DEL DNA CON PROTEÍNAS

Son numerosos los procesos biológicos en los que la función del DNA depende de su interacción específica con proteínas; destaca entre ellos el control de la expresión génica, ejercido por proteínas denominadas genéricamente *factores de transcripción* (cuyo papel se estudia más adelante, pág. 264). El estudio de estas **proteínas** tiene interés dentro del contexto de las variantes estructurales del DNA, precisamente por su capacidad de **reconocer e interactuar con secuencias concretas del DNA**. Este reconocimiento tiene lugar gracias a la posibilidad de variaciones locales en la estructura tridimensional del DNA, así como a la interacción directa de la proteína con las bases, sin necesidad de que se separen las hebras de la doble hélice. En la mayor parte de los casos, los numerosos contactos establecidos (de tipo iónico, hidrofóbico y enlaces de hidrógeno), aunque individualmente débiles, se combinan para dar una interacción específica y muy fuerte entre DNA y proteína.

El reconocimiento se ejerce a través de regiones o elementos estructurales, que son sólo una parte de la estructura terciaria completa de la proteína, y que se pueden encontrar, con una estructura espacial muy similar, en proteínas distintas, interviniendo en una misma función en todas ellas. Si se cumplen estos requisitos, reciben la denominación genérica de "motivos" estructurales proteicos (en inglés, *motifs*), en este caso "motivos de unión a DNA".

La interacción tiene lugar con el exterior de la doble hélice, especialmente sobre el surco mayor. Como característica importante, un mismo motivo puede contactar simultáneamente con varios puntos diferentes del DNA, forzando así en éste una modificación conformacional que hace que secuencias lejanas puedan aproximarse. Dicho cambio conformacional puede inducir incluso la separación parcial de las hebras (desnaturalización local), con lo que se vería favorecido el acceso de otras proteínas. Tal es el caso, por ejemplo, de la RNA-polimerasa en el proceso de transcripción.

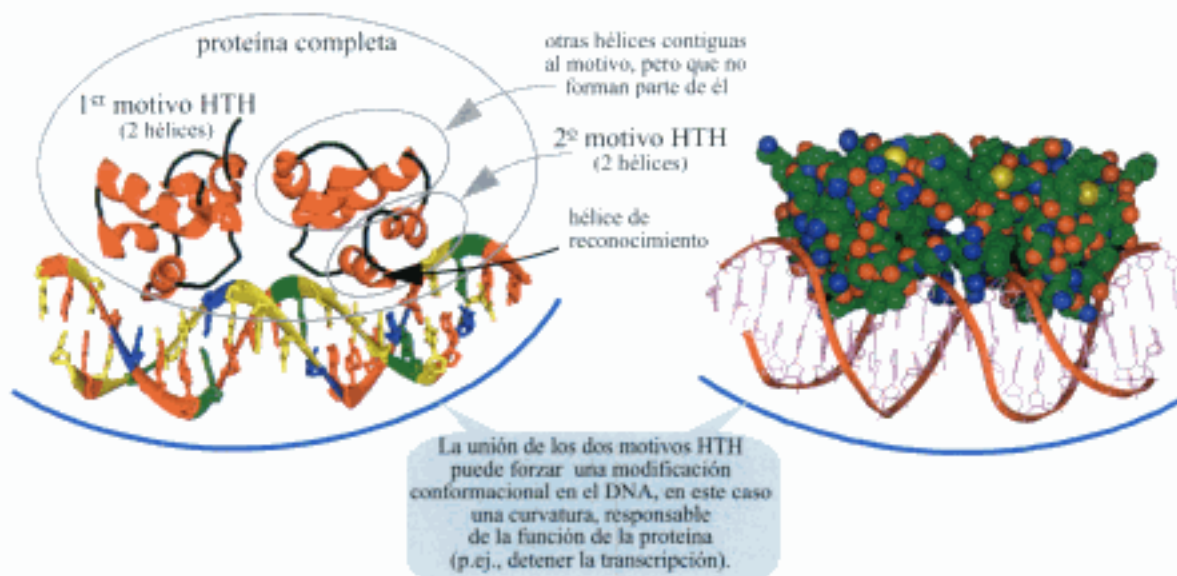
6.3.1 Descripción de los motivos estructurales más frecuentes



6.3.1.1 Motivo hélice-giro-hélice o hélice-vuelta-hélice

Se abrevia como **HTH**, del inglés *helix-turn-helix*. Fue el primer motivo proteico de unión al DNA bien estudiado. Se ha encontrado en proteínas reguladoras de la expresión génica, tanto en virus como en procariotas y eucariotas. Está formado por dos segmentos peptídicos en α -hélice, de estructura rígida, separados por una secuencia de aminoácidos, flexible, que permite que las dos hélices se aproximen entre sí. Una de las α -hélices (llamada "hélice de reconocimiento") se encaja en el surco mayor del DNA, de forma que los residuos aminoácidos de un lado de la hélice interaccionan mediante puentes de hidrógeno con las bases nitrogenadas expuestas en ese surco. El diámetro de la α -hélice tiene la dimensión adecuada (1,2 nm) para entrar en el surco. Algunos aminoácidos, situados en la cara opuesta de la misma hélice, establecen interacciones con otros de la segunda hélice, fijando la posición relativa de ambas en un ángulo casi perpendicular (ver esquema inicial). La segunda α -hélice queda así cruzada sobre la de reconocimiento y situada en el exterior del DNA.

Dos modelos moleculares donde se aprecia la interacción del DNA con un motivo hélice-giro-hélice dimérico. En este caso, la proteína completa contiene 2 motivos HTH. Se muestran sólo éstos y algunas hélices cercanas, ajenas al motivo; el resto de la molécula no se muestra, pues no se relaciona con la unión al DNA.

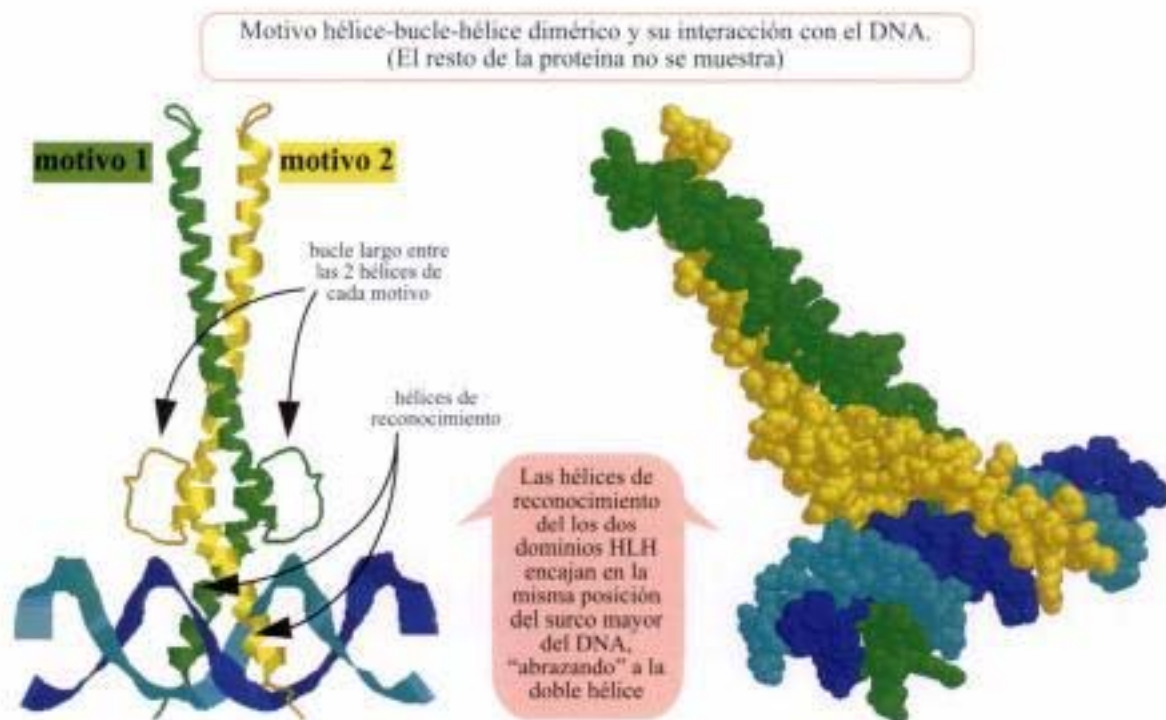


Es frecuente que una proteína presente dos motivos HTH o bien que se asocien dos moléculas de proteína iguales (es decir, un dímero), cada una con un motivo HTH. En ambos casos, las dos hélices de reconocimiento (una de cada motivo) se sitúan a una distancia mutua de unos 3,4 nm, lo que coincide con el paso de hélice del B-DNA (3,54 nm) y, por tanto, con dos posiciones consecutivas de su surco mayor. Es decir, ambos motivos interaccionan simultáneamente con el DNA, contribuyendo a estabilizar la unión de la proteína.

Los ejemplos mejor conocidos de proteínas con motivos HTH actúan regulando la transcripción en bacterias y virus, respectivamente. El primero, la *proteína receptora de AMP cíclico (CRP)* (también llamada CAP) es un dímero, por tanto, con 2 motivos HTH, que cuando un cAMP aumenta su afinidad por secuencias específicas del DNA, desencadenando la expresión de los genes del operón lactosa (este fue el primer ejemplo conocido de control de la transcripción). El segundo ejemplo (en la figura) es la proteína **Cro** del fago lambda, codificada por un gen del virus y expresada en la bacteria anfitriona, que interacciona por su motivo HTH con el gen de la proteína vírica llamada *represor del fago λ* , bloqueando como consecuencia su expresión (transcripción del gen y síntesis de la proteína).

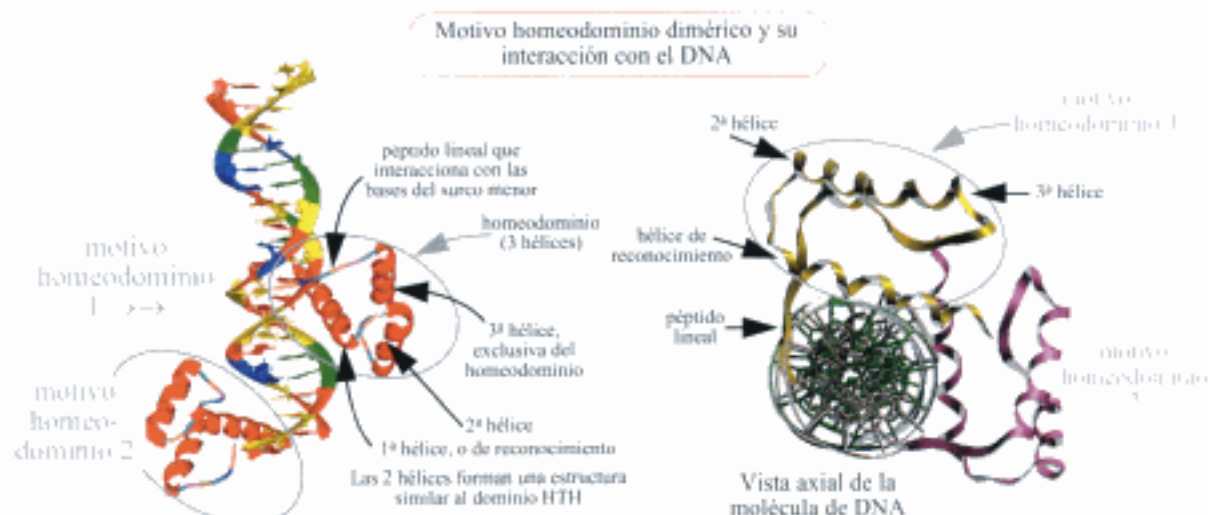
6.3.1.2 Motivo hélice-bucle-hélice

Se abrevia como **HLH**, del inglés *helix-loop-helix*. No debe confundirse con el anterior. Consta también de dos segmentos en α -hélice, pero en este caso el péptido que los une es más largo, lo que lo dota de mayor flexibilidad y, en consecuencia, hay más posibilidades de orientación mutua de las dos hélices. Generalmente, al igual que en HTH, dos motivos HLH se asocian, formándose proteínas dimericas, aunque no siempre interaccionan en posiciones consecutivas del surco mayor. Un ejemplo de motivo HLH bien conocido es el de la proteína **Max** de ratón (en la figura).



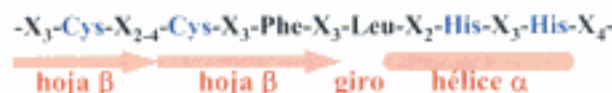
6.3.1.3 Motivo homeodominio

Se puede considerar como una ampliación del motivo hélice-giro-hélice (HTH), pero adquiere entidad propia por aparecer repetidamente, con idéntica estructura, en distintas proteínas. Además de la disposición de dos hélices cruzadas típica del HTH, existe un tercer tramo en α -hélice, del que sale una región sin estructura secundaria definida cuyos aminoácidos interaccionan directamente con el DNA. Este motivo (en inglés, *homeodomain*) es importante por aparecer en las proteínas que regulan el desarrollo embrionario, estudiadas especialmente en *Drosophila*.

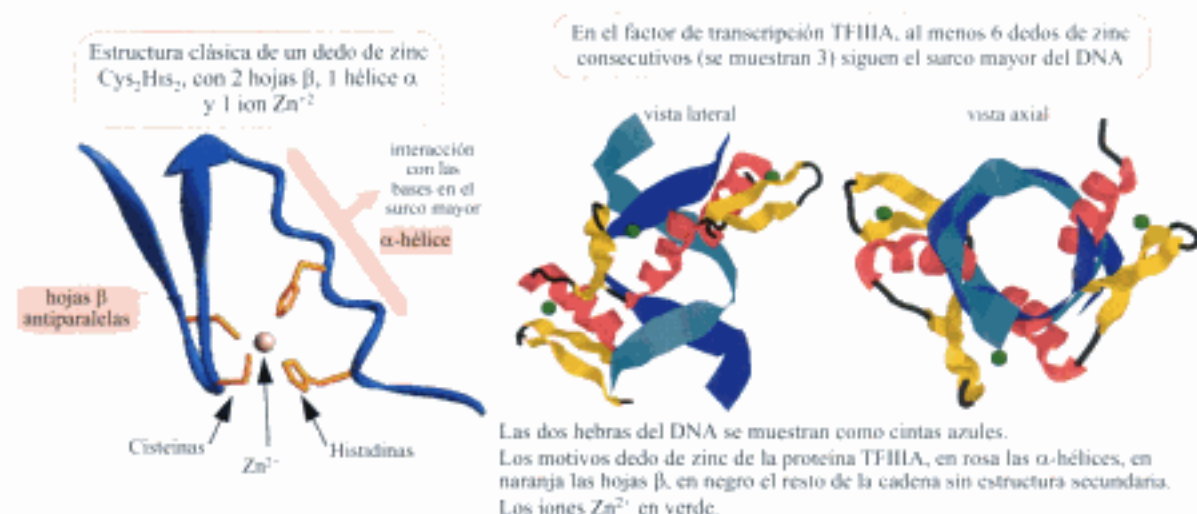


6.3.1.4 Motivo dedo de zinc

Se ha encontrado este tipo de motivo estructural (en inglés, *zinc finger*) en multitud de proteínas eucarióticas que se unen al DNA, entre ellas el factor de transcripción TFIIIA y el receptor de estrógenos. Al igual que en las estructuras anteriores, es frecuente que una misma proteína posea más de un motivo; en este caso, se observan normalmente múltiples dedos de zinc consecutivos. Éste es un elemento estructural formado por unos 30 aminoácidos, de los cuales 2 cisteínas y 2 histidinas aparecen en posiciones constantes, coordinando tetraédricamente un ion Zn^{2+} . Esta unión obliga a un plegamiento de la cadena peptídica, de forma que el motivo se caracteriza por una conformación tridimensional alargada, en forma de "dedo", que le da nombre. La variante más frecuente muestra dos hojas β antiparalelas, cada una con un residuo Cys, seguidas de un giro y de una estructura α -helicoidal, con las dos His. Este motivo se denomina C_2H_2 (abreviaturas de una letra para los aminoácidos que unen el Zn^{2+}) y su secuencia consenso es:



donde los subíndices representan el número de residuos (aminoácidos variables, X) que separa los aminoácidos consenso (constantes). Estas secuencias adquieren espontáneamente gracias al Zn^{2+} la conformación apropiada para que el segmento α -helicoidal interactúe con tres bases expuestas en el surco mayor del DNA. Los restos variables permiten que se reconozcan distintas secuencias del DNA.



En otros casos se observa una variante de dedo de zinc llamada C_4 , donde el ion metálico se coordina con cuatro Cys en la secuencia consenso

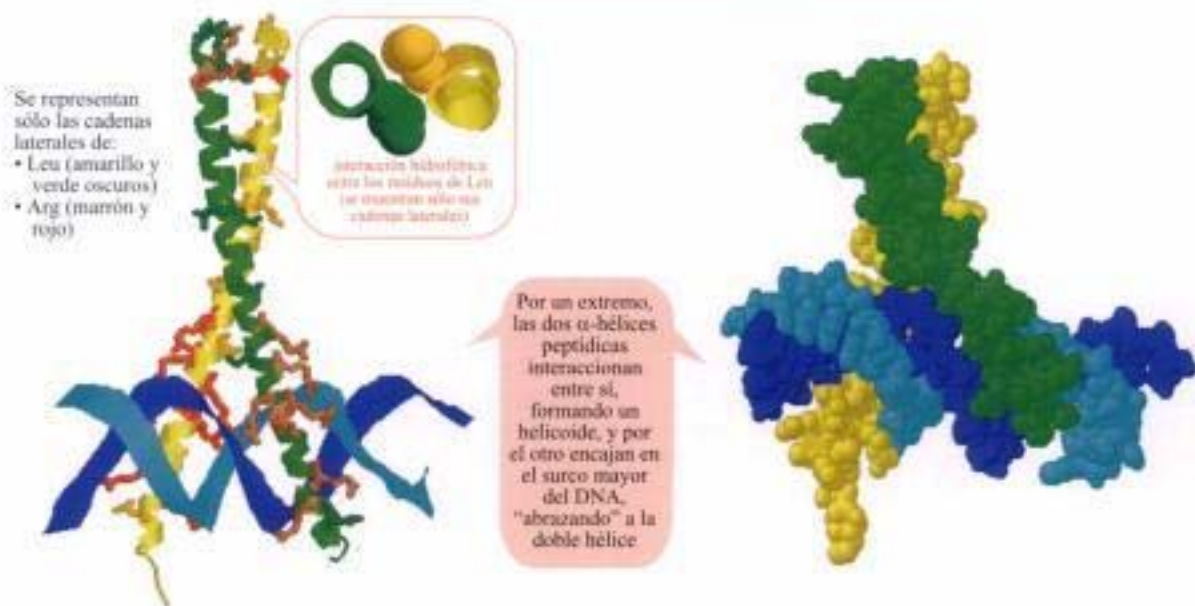


que no forma hojas β , sino dos segmentos en α -hélice. Esta estructura aparece, por ejemplo, en los receptores de hormonas esteroideas.

6.3.1.5 Motivo cremallera de leucina

Se trata de una región de la proteína cuya secuencia tiene un residuo de leucina cada 7 aminoácidos. Al adoptar la conformación α -helicoidal se concentran en un lado los aminoácidos hidrofóbicos, de tal manera que Leu se repite en la misma cara cada dos vueltas de hélice (2 vueltas de 3,4 residuos = 7 aminoácidos). Dos cadenas peptídicas de este tipo pueden asociarse hidrofóticamente, intercalando sus restos de Leu como si se tratara de los dientes de una cremallera (en inglés, *leucine zipper*). En el otro extremo del motivo las dos α -hélices encajan en el surco mayor del DNA y presentan en la cara exterior abundantes aminoácidos básicos que interaccionan favorablemente con los fosfatos. De alguna manera, este modelo se asemeja a una "Y", en la que el tallo correspondería a la cremallera de leucina y los brazos a los tramos de las hélices que contactan con el DNA. Estructuras de este tipo aparecen, por ejemplo, en las proteínas reguladoras de la transcripción llamadas bZIP.

Motivo cremallera de leucina y su interacción con el DNA



Tema 7

Estructuras de orden superior de DNA y RNA

7.1	Superenrollamiento del DNA	66
7.1.1	Superenrollamiento del DNA en procariotas (y mitocondrias y cloroplastos de eucariotas)	66
7.1.1.1	Conformación relajada	66
7.1.1.2	Conformaciones superenrolladas	67
7.1.1.3	Demostración experimental del superenrollamiento	67
7.1.2	Aspectos teóricos del superenrollamiento	68
7.2	Estructura de los RNAs	69
7.2.1	Composición de bases y estructuras primaria y secundaria del RNA	69
7.2.2	Estructura tridimensional de los RNAs en general	70
7.2.3	Tipos de RNA: propiedades y estructuras particulares	70
7.2.3.1	RNA mensajero (mRNA)	71
7.2.3.2	RNA de transferencia (tRNA)	71
a)	Especificidad de los tRNAs por los aminoácidos	71
b)	Presencia de nucleósidos infrecuentes	72
c)	Apareamiento de bases intracatenario	73
d)	Estructura secundaria en trébol	74
e)	Estructura terciaria en "L"	75
7.2.3.3	RNA ribosómico (rRNA)	76
7.2.4	Ribozimas: función catalítica de algunos RNAs	76
7.2.4.1	Acción catalítica en la maduración de RNA	77
7.2.4.2	Acción autocatalítica en la eliminación de intrones	77
7.2.4.3	Acción catalítica y autocatalítica en viroides: ribozimas de posible aplicación terapéutica	78
7.2.4.4	Acción catalítica transpeptidasa en la síntesis de proteínas	78
7.3	Los ribosomas	79
7.3.1	Localización celular	79
7.3.2	Características diferenciales de los ribosomas procarióticos y eucarióticos	80
7.3.2.1	Origen	80
7.3.2.2	Composición	80
7.3.2.3	Estructura	81
7.3.3	Separación de los ribosomas, subunidades y componentes	81

Los ácidos nucleicos no se encuentran en las células en las formas extendidas correspondientes a su estructura primaria y secundaria, sino molecularmente más compactados, lo que hemos denominado (p. 32) nivel estructural "de orden superior". En el caso del DNA, parte de esta compactación o condensación viene representada por el denominado *superenrollamiento*, o retorcimiento de la cadena sobre sí misma, mientras que para el RNA se plasma en estructuras tridimensionales variadas. En ambos casos, la condensación se completa por la asociación estrecha con proteínas que inducen el plegamiento de la doble hélice del DNA o de la cadena lineal de algunos RNAs (p.ej., el RNA ribosómico, pág. 79). La contribución de esta asociación con proteínas posee una relevancia especial en el caso del DNA nuclear eucariótico (tema 8).

7.1 SUPERENROLLAMIENTO DEL DNA

El superenrollamiento, como parte de la condensación del DNA en general, posee especial significado para entender la organización del cromosoma. En concreto, contribuye a explicar cómo un DNA de tan gran magnitud puede alojarse en el interior de la célula (procariotas) o del núcleo (eucariotas), de dimensiones muy inferiores a la longitud teórica del DNA en doble hélice. Por ejemplo:

DNA	tamaño	longitud en forma de B-DNA	dimensiones en las que se compacta
<i>E. coli</i>	$4,7 \cdot 10^6$ pb	1,6 mm = 1.600 μ m	2 μ m (célula)
Cada cromosoma humano	$50 - 250 \cdot 10^6$ pb	17 - 85 mm	4-6 μ m (núcleo)
Genoma humano diploide	$6,6 \cdot 10^9$ pb	2,2 m	

Se entiende por **superenrollamiento** (enrollar algo que ya está enrollado) del DNA al retorcimiento o giro sobre sí misma de la doble hélice (que ya lleva implícito el **enrollamiento** mutuo de dos hebras), de forma que el **eje** de la doble hélice no sigue una línea recta, sino otra hélice (una "superhélice"). Su causa, como se verá a continuación, es la introducción de una tensión estructural en la propia molécula. Aparece de algún modo en todos los DNAs celulares (ya enrollados) y se encuentra estrechamente regulado *in vivo* por la acción de las topoisomerasas (pág. 151). En resumen, las dos hebras del DNA sufren un enrollamiento, mientras que la doble hélice sufre un superenrollamiento.



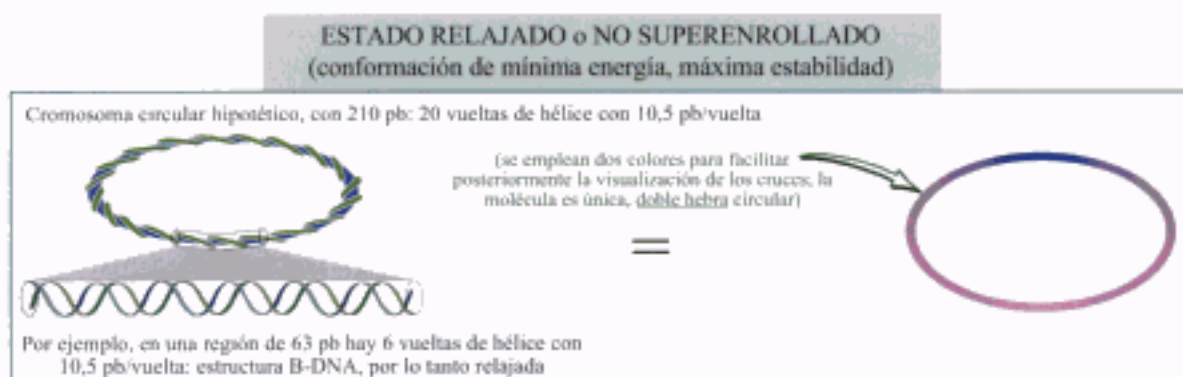
Un ejemplo intuitivo del concepto de superenrollamiento es el cordón del teléfono, que de por sí presenta un enrollamiento helicoidal y, con el uso cotidiano, termina por adoptar un enrollamiento adicional (cordón retorcido). Este símil se utilizó para explicar muchas propiedades del DNA vírico (circular y pequeño). Un modelo más próximo a la estructura del DNA es una cuerda formada por dos cabos arrollados entre sí. Aunque la cuerda tiende a mantener ese arrollamiento, puede ser forzada manualmente a un mayor o menor número de vueltas entre sus cabos (torsión), con lo que se genera una tensión que se libera mediante la formación de superenrollamientos.

7.1.1 Superenrollamiento del DNA en procariotas (y mitocondrias y cloroplastos de eucariotas)

Aunque afecta tanto a procariotas como a eucariotas, el superenrollamiento es más conocido y de comprensión más sencilla en los primeros, por lo que así se abordará a continuación. Estas consideraciones son igualmente válidas para el DNA de plásmidos procarióticos y el de cromosomas mitocondriales y cloroplásticos de eucariotas. En todos estos casos, la molécula de DNA muestra una estructura circular o cerrada (una doble hélice con sus extremos unidos entre sí), en la que es fácil apreciar los fenómenos de superenrollamiento.

7.1.1.1 Conformación relajada

La conformación adoptada por el DNA en su forma B (aprox. 10,5 pb/vuelta) corresponde a un estado **estable**, de mínima energía, que se considera **no superenrollado** o **relajado**, cuando el eje de la doble hélice se puede disponer completamente sobre un plano.



7.1.1.2 Conformaciones superenrolladas

En las células, el DNA generalmente no se encuentra en la conformación relajada, sino en un estado **no relajado**, **superenrollado** y más compacto. Este es producido por topoisomerasas que, mediante el corte transitorio de una o ambas hebras, introducen un aumento o una disminución en el número de vueltas de hélice (pág. 151). En ambos casos se produce una tensión estructural, debida a un plegamiento local de las hebras y a una posición relativa de los nucleótidos diferentes de su geometría energéticamente más estable, la del B-DNA. Esta tensión se contrarresta en la molécula mediante la formación de un superenrollamiento: la cadena entera del DNA gira sobre sí misma de forma que las bases puedan disponerse de nuevo del modo más próximo a la conformación B-DNA.

En primer lugar, si se **reduce** el número de vueltas de hélice, la tensión se libera formando una superhélice a la que se le asigna un valor **negativo** de superenrollamiento.

ESTADO NO RELAJADO o TENSO (de energía superior) originado por **desenrollamiento** de la doble hélice (pérdida de vueltas)

Conformación de mínima energía, obtenida mediante un **SUPERENROLLAMIENTO NEGATIVO**

La molécula de DNA ha sufrido una **pérdida** de una vuelta:

El cromosoma de 210 pb contiene ahora 19 vueltas de hélice, con 11,1 pb/vuelta



la **tensión** se libera (pérdida de energía) mediante la formación de un **bucle** o **superhélice**

compactación de la molécula de DNA circular por un superenrollamiento **negativo**



Si se considera que la pérdida de una vuelta de hélice ha ocurrido en la región de 63 pb, ésta tendrá 5 vueltas de hélice con 12,6 pb/vuelta: > 10,5 pb/vuelta, es decir, **tensa**

El cromosoma superenrollado tiene 20 vueltas con 10,5 pb/vuelta: conformación **no tensa**

Si, por el contrario, se **aumenta** el número de vueltas de hélice, la tensión se libera por formación de una superhélice de sentido opuesto, con un superenrollamiento **positivo**.

ESTADO NO RELAJADO o TENSO (de energía superior) originado por **superenrollamiento** de la doble hélice (aumento de vueltas)

Conformación de mínima energía, obtenida mediante un **SUPERENROLLAMIENTO POSITIVO**

La molécula de DNA ha sufrido un **aumento** de una vuelta:

El cromosoma de 210 pb contiene ahora 21 vueltas de hélice, con 10,0 pb/vuelta



la **tensión** se libera (pérdida de energía) mediante la formación de un **bucle** o **superhélice** (de sentido contrario al anterior)

compactación de la molécula de DNA circular por un superenrollamiento **positivo**



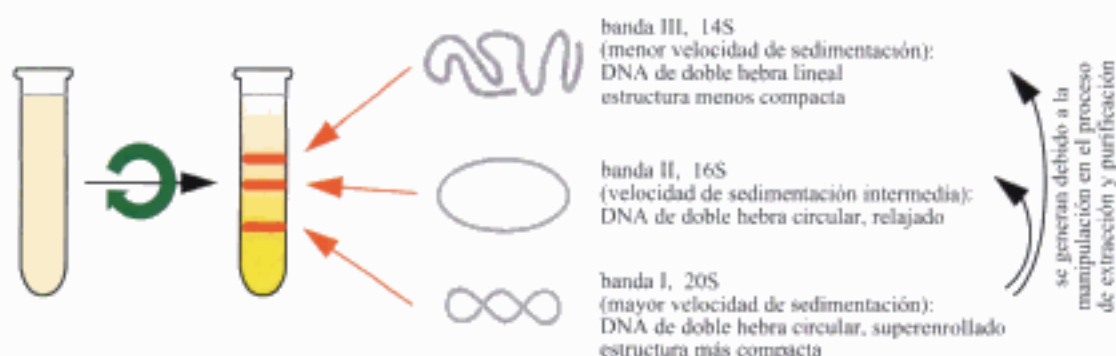
Si se considera que el aumento de una vuelta de hélice ha ocurrido en la región de 63 pb, ésta tendrá 7 vueltas de hélice con 9,0 pb/vuelta: < 10,5 pb/vuelta, por tanto **tensa**

El cromosoma superenrollado tiene 20 vueltas con 10,5 pb/vuelta: conformación **no tensa**

7.1.1.3 Demostración experimental del superenrollamiento

De estas dos posibilidades, el DNA se encuentra normalmente en las células en un estado de **superenrollamiento negativo**. La primera evidencia experimental que llevó a proponer la hipótesis del superenrollamiento de las cadenas se obtuvo gracias a la aplicación de la técnica de **ultracentrifugación isopícnica** (basada en el equilibrio de sedimentación en un gradiente de densidad de cloruro de cesio, pág. 137).

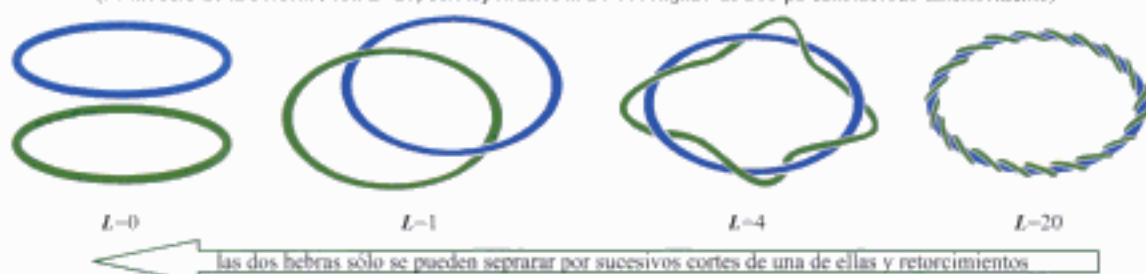
Se observó que el DNA de *polioma*, un pequeño virus que causa cáncer a ratones, se separaba en tres bandas en el gradiente. Las tres fracciones, sin embargo, tenían la misma masa molecular, por lo que la diferencia debía estar en su forma o estructura. Se pudo comprobar que las moléculas menos densas eran lineales, mientras que las demás eran circulares. Sólo tras la hipótesis del superenrollamiento se pudo explicar la diferencia entre esas dos fracciones más densas.



7.1.2 Aspectos teóricos del superenrollamiento

El empleo de la **Topología** (rama de las Matemáticas que estudia las propiedades de posición relativa de las partes de un objeto, propiedades que no cambian al ser sometido a deformación) ha permitido aclarar la estructura del DNA superenrollado. El signo y magnitud del superenrollamiento se asignan con base en dicho estudio topológico, del cual sólo se comentan aquí los aspectos más elementales.

Posibles disposiciones teóricas de dos hebras (por ejemplo, dos cabos de cuerda) con diferente grado de enrollamiento entre sí (diferente número de enlace). L = número de enlace = nº de veces que las hebras se cruzan.
(El modelo de la derecha, con $L=20$, correspondería al DNA relajado de 210 pb considerado anteriormente)



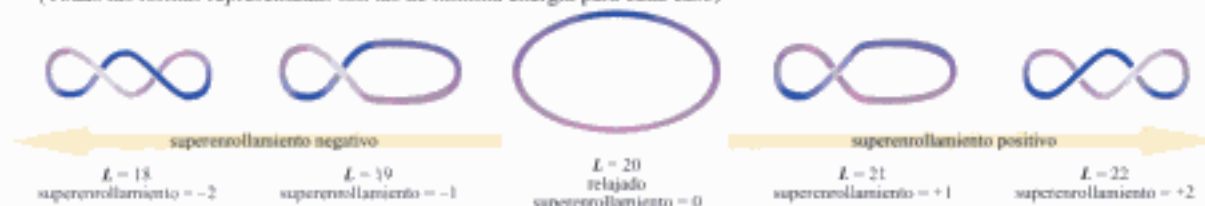
Se define el **índice o número de enlace topológico** (L o Lk , de *linking number*) como el número de veces que las dos hebras se cruzan entre sí; más específicamente, el número de veces que una hebra cruza el plano definido por la otra (como anillo o curva cerrada). Se le asigna valor positivo cuando las hebras se cruzan hacia la derecha.

Se habla de **enlace topológico** porque cuando $L \neq 0$ las dos hebras están físicamente ligadas (es decir, no se pueden separar), a pesar de que no hay enlaces covalentes entre ellas. Sólo se podrían separar ($L=0$) rompiendo algún enlace covalente (es decir, cortando una hebra). Lo mismo es aplicable para cualquier cambio del valor L . Por la misma razón, la deformación, plegamiento, etc., de las hebras, sin cortarlas, no alteran el número de enlace ni el número de superenrollamiento.

Para el caso del DNA, la conformación relajada posee, por tanto, un número de enlace igual al número de vueltas de hélice, siempre positivo puesto que la doble hélice del DNA es dextrógira. Como ya se ha estudiado, las situaciones de superenrollamiento surgen de cambios en el número de vueltas de hélice, es decir, en el número de enlace L .

Para discernir y cuantificar los estados superenrollados se refiere su valor de L al del estado relajado, definiendo un nuevo valor numérico de **superenrollamiento** como la diferencia (ΔL) entre el número de enlace de una conformación concreta del DNA y el número de enlace de su forma relajada. Este valor de superenrollamiento puede, lógicamente, ser positivo o negativo.

Para un DNA doble circular de 210 pb de longitud, la estructura B tiene $210 \text{ pb} / 10,5 \text{ pb/vuelta} = 20$ vueltas. Por tanto, la conformación relajada de ese DNA sería la de $L=20$.
Por encima y por debajo de $L=20$ se habla de un superenrollamiento positivo o negativo, respectivamente.
(Todas las formas representadas son las de mínima energía para cada caso)



7.2 ESTRUCTURA DE LOS RNAs

Como ya se ha indicado (pág. 33), los RNAs son polirribonucleótidos (polinucleótidos con ribosa, y U en lugar de T) formados por cadenas lineales, de hebra sencilla, de varias decenas o millares de unidades, pero de longitud muy inferior a las de DNA. No se presentan como cadenas dobles de hebras complementarias.

7.2.1 Composición de bases y estructuras primaria y secundaria del RNA

La estructura primaria de los RNAs ha sido estudiada anteriormente (tema 4). En cuanto a la secundaria, es característica la ausencia de una estructura repetitiva y bien definida, a diferencia del DNA; realmente sólo existe estructura secundaria en parte de la molécula de tRNAs y rRNAs.

Al no existir un apareamiento sistemático de dos hebras, la composición de bases en los RNAs es mucho más variada que en el DNA y no se cumplen las reglas de Chargaff (pág. 38). Por otro lado, en algunos RNAs hay cierta abundancia de nucleótidos distintos de los cuatro principales (A, G, C y U), aunque en proporción baja con respecto a la de éstos.

Composición de bases de varios RNAs (%)

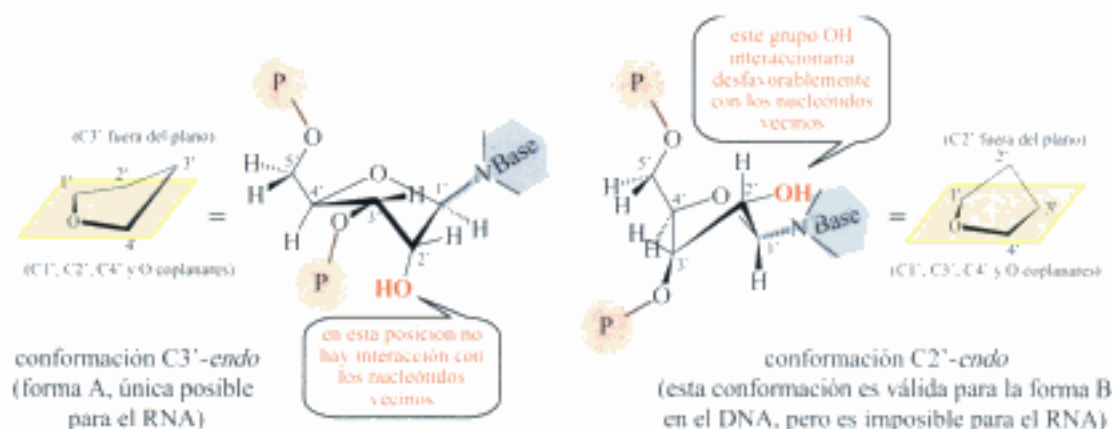
Organismo	Tipo de RNA	A	G	C	U
Rata (hígado)	nuclear	20.2	25.7	29.5	24.6
	mitocondrial	17.8	31.8	28.4	20.9
	ribosómico	20.0	30.5	31.6	20.2
	de transferencia	20.9	30.9	28.5	20.4
	mensajero	23.8	28.3	27.4	20.5
Levaduras	ribosómico	24.9	27.7	19.4	26.7
	de transferencia	18.5	29.2	28.4	20.0
	mensajero	27.5	25.1	20.3	27.1
<i>E. coli</i>	ribosómico	~ 25	~ 31	~ 22	~ 21
	de transferencia	19.3	32.0	28.3	16.0
	mensajero	24.1	27.7	24.7	23.5

Obsérvese cómo, a diferencia de lo que ocurre con el DNA,

- la composición de bases varía dentro de una misma especie (según el RNA considerado)
- no existe una equivalencia entre bases ($A \neq U$, $G \neq C$, purinas $\neq$ pirimidinas)

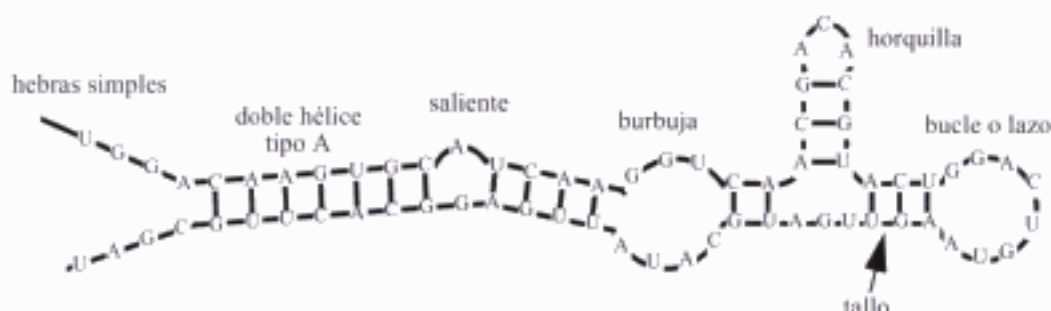
Algunos RNAs adoptan estructura secundaria en pequeñas regiones de su molécula, gracias al plegamiento de su única hebra sobre sí misma, siempre que éste permita el apareamiento *intracatenario* de dos zonas cuyas secuencias sean complementarias, a pesar de estar separadas en la estructura primaria; se forma así una doble hélice local, de carácter antiparalelo. De manera similar se forman también dobles hélices entre una molécula de RNA y una hebra sencilla de DNA (híbridos RNA-DNA).

En estos casos, si el grupo 2'-OH de la ribosa adopta la conformación C2'-endo presenta impedimentos estéricos con otros átomos de la cadena de RNA, lo que impide que se forme la doble hélice (que sería de tipo B). En concreto, el 2'-OH estaría muy cerca de los átomos del grupo fosfato contiguo y del C8 de la base adyacente. En cambio, en la conformación C3'-endo el grupo 2'-OH se proyecta hacia fuera, lejos de otros átomos del esqueleto y no aparece impedimento estérico alguno, por lo que es posible la formación de la doble hélice, en este caso de tipo A. En resumen, la única doble hélice en la que puede participar un RNA es la de tipo A (el DNA, al carecer del OH en 2', puede adoptar tanto la A como la B, como ya se ha estudiado en el tema 6).



7.2.2 Estructura tridimensional de los RNAs en general

Existe una gran diversidad en la estructura tridimensional de las moléculas de RNA. Ésta es compleja, única para cada tipo de RNA, y resulta de la combinación de estructuras secundarias locales, estabilizadas por enlaces de hidrógeno y por interacciones hidrofóbicas de apilamiento de bases. Su estudio concreto se hará más adelante, para cada tipo de RNA.



A pesar de la inexistencia de una estructura secundaria propiamente dicha para toda la molécula, la cadena sencilla de cualquier RNA adopta localmente variadas estructuras: horquillas, cruciformes, bucles, lazos, etc. Algunas de ellas se deben al apareamiento de bases *intracatenario* formando cortos tramos de doble hélice de tipo A, como se acaba de indicar. La proporción de zonas helicoidales varía ampliamente entre distintos RNAs, pero puede alcanzar hasta un 50% de la estructura total.

7.2.3 Tipos de RNA: propiedades y estructuras particulares

Tanto en procariotas como en eucariotas existen tres tipos principales de RNA: mensajero (mRNA), de transferencia (tRNA) y ribosómico (rRNA). Además de ellos, en eucariotas existen RNA nuclear heterogéneo (hnRNA), nuclear pequeño (snRNA) y citoplásmico pequeño (scRNA), y RNAs de orgánulos (mitocondrias y cloroplastos). Todos ellos tienen en común una función relacionada, directa o indirectamente, con la expresión génica (maduración postranscripcional y traducción). Adicionalmente, en algunos virus el RNA juega el papel de portador de la información genética, que en el resto de organismos corresponde al DNA.

Las diferencias fundamentales entre los distintos tipos de RNA se resumen en la tabla.

Clase	Ubicación celular	Cantidad	Tamaño		Características particulares
			s	nº nucleótidos	
mRNA	citoplasma (P y E)	5%		600-3.000	estructura sencilla lineal
tRNA	citoplasma (P y E)	20%		75 a 95	hay 50-60 diferentes, específicos para cada aa
rRNA	citoplasma (P y E)	75%			
			3 tipos, en ribosomas de procariotas:		
			16S	1.500	forma parte de la subunidad pequeña del ribosoma
	23S		2.900	} forman parte de la subunidad grande del ribosoma	
	5S		120		
	4 tipos, en ribosomas de eucariotas:		18S	1.900	forma parte de la subunidad pequeña del ribosoma
			28S	4.700	} forman parte de la subunidad grande del ribosoma
			5,8S	160	
5S		120			
hnRNA		núcleo (E)	transitorio	200-30.000	transcrito primario, precursor del mRNA
snRNA	núcleo (E)	transitorio	100-300	forma parte de ribonucleoproteínas nucleares pequeñas (snRNPs)	
scRNA	citoplasma (E)	transitorio			
mtRNA	mitocondrias (E)	transitorio			
E = eucariotas P = procariotas					
s = coeficiente de sedimentación (medido en Svedbergs, pág. 123)					

7.2.3.1 RNA mensajero (mRNA)

La propuesta de la existencia de una molécula que transfiere el mensaje genético desde el núcleo al citoplasma se debe a François Jacob y Jacques Monod (premios Nobel en 1965). Dicha molécula es el RNA mensajero que, al ser una copia de la información contenida en la secuencia del DNA, actúa posteriormente en el ribosoma como molde o plantilla para la síntesis proteica.

Las características principales del mRNA son las siguientes:

- Se encuentra en el citoplasma de todas las células, en baja proporción respecto al RNA total y con una vida media corta dada su rápida síntesis y degradación. Aparece asociado transitoriamente al ribosoma.
- Sus moléculas son de tamaño muy diverso (en relación con el tamaño de las proteínas que codifican).
- Es el tipo de RNA de estructura más sencilla, con una cadena lineal, formada exclusivamente por los ribonucleósidos A, U, C y G. La composición de bases es un reflejo de la del DNA molde del que proceden. La molécula carece de una conformación definida: la cadena se mantiene extendida, sin adoptar estructuras secundarias ni terciarias significativas, gracias a la participación de proteínas que se asocian a ella.

7.2.3.2 RNA de transferencia (tRNA)

Fue Francis Crick quien sugirió que los aminoácidos se unían a una molécula intermedia, o "adaptador", antes de interaccionar con el mRNA. Efectivamente, los tRNAs actúan como moléculas adaptadoras en la síntesis de proteínas. Esta función depende de una doble interacción:

- Por un extremo de la molécula unen el aminoácido para dar aminoacil-tRNA (pág. 302), la forma activa necesaria para la reacción de formación del enlace peptídico.
- Por el otro extremo se aparean con el mRNA, en presencia del ribosoma, transfiriendo el aminoácido al péptido creciente. De esta manera, los aminoácidos se ordenan correctamente en la proteína, de acuerdo con la secuencia leída por el tRNA sobre el mRNA (pág. 311).



A pesar de ser los RNAs de menor tamaño, los tRNAs presentan una gran complejidad estructural. Se han aislado y estudiado muchos, de distintos organismos, mostrando una estructura común a pesar de sus distintas secuencias, que reflejan la necesidad de adaptación a su especificidad.

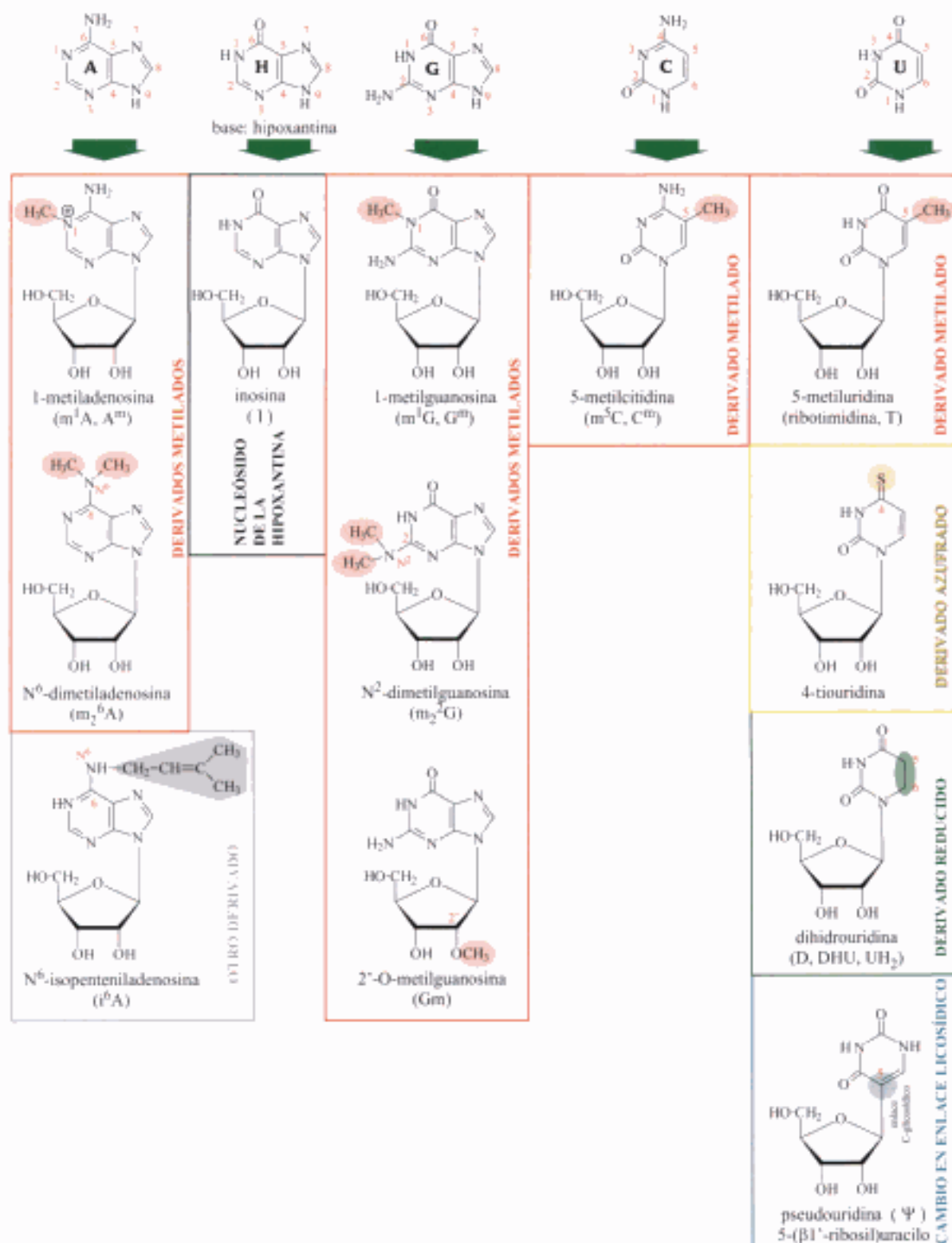
a) Especificidad de los tRNAs por los aminoácidos

Los tRNAs constituyen una fracción importante del RNA total y se encuentran disueltos en el citosol. Una célula puede contener hasta 60 tRNAs diferentes, es decir, en número superior al de aminoácidos (20), puesto que cada tRNA interacciona específicamente con un solo aminoácido, pero algunos aminoácidos son reconocidos por más de un tRNA (se habla en este caso de tRNAs "sinónimos", pág. 296).

La especificidad se indica con el nombre del aminoácido en superíndice (por ej., tRNA^{Ala}); los tRNAs sinónimos llevan subíndices numéricos (por ej., tRNA^{Trp1} y tRNA^{Trp2}). Cuando un tRNA está unido con el aminoácido se dice que está "activado" o "cargado", se denomina aminoacil-tRNA y se indica con el aminoácido como prefijo (por ejemplo, Ala-tRNA^{Ala}, Trp-tRNA^{Trp1}, Trp-tRNA^{Trp2}, etc.).

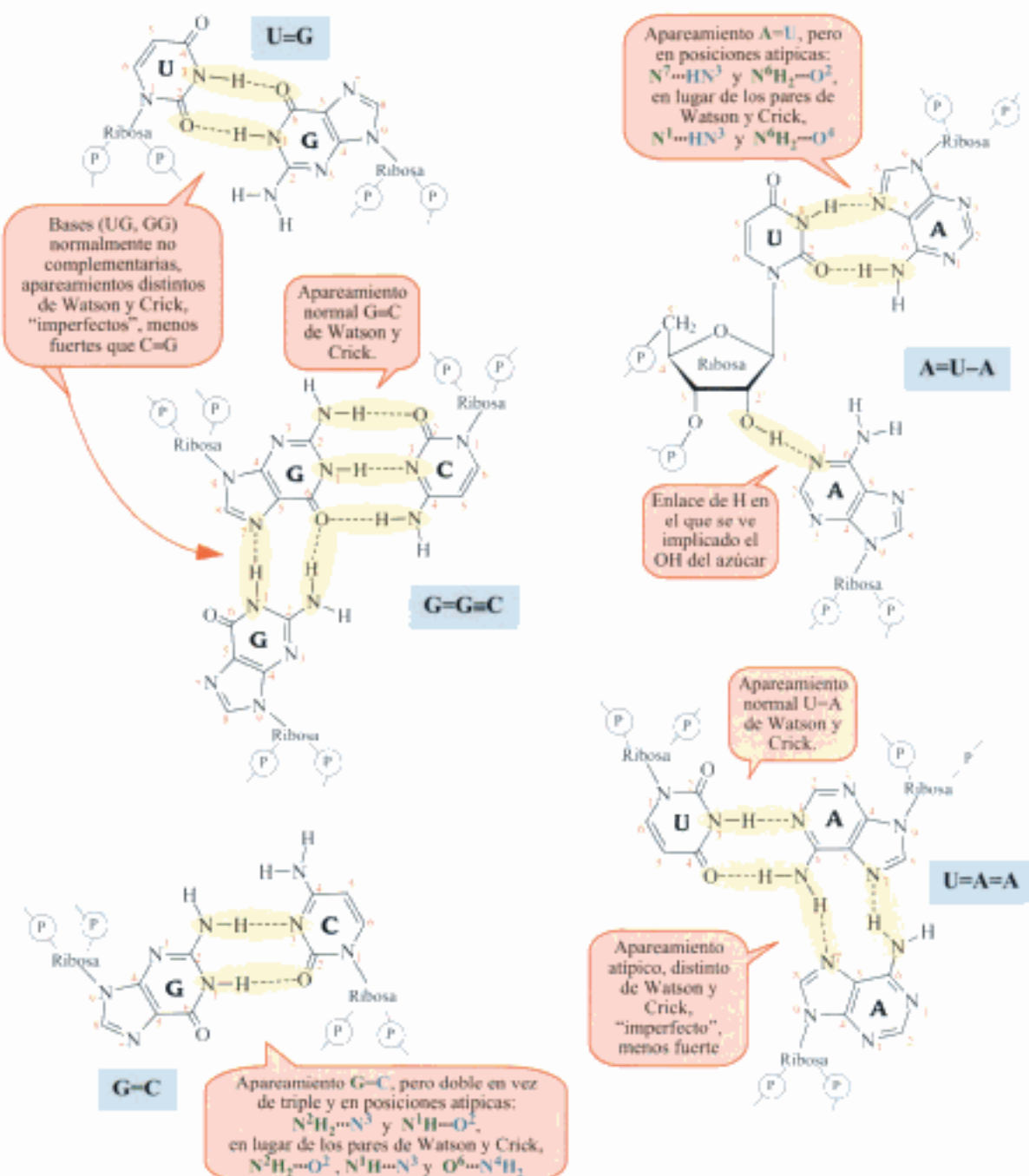
b) Presencia de nucleósidos infrecuentes

Una de las características de los tRNAs es que en su composición presentan, además de los cuatro principales, otros nucleósidos no corrientes o **infrecuentes**, que pueden suponer hasta un 10% del total. Generalmente se trata de derivados de los cuatro principales, A, G, C, U, así como de la hipoxantina, a menudo con grupos metilo en la base o la desoxirribosa. Además, es característica la presencia de ribotimidina. La presencia de todos estos nucleósidos está relacionada con la estructura terciaria y tiene por ello una clara influencia sobre la función.



c) *Apareamiento de bases intracatenario*

A pesar de la abundancia de nucleótidos no usuales, los tRNAs muestran un cierto grado de complementariedad, es decir, de **regiones con estructura secundaria en doble hélice** basada en apareamientos de bases **intracatenarios**. Con frecuencia se observan apareamientos distintos de los de Watson y Crick, llamados genéricamente **de Hoogsteen**: por ejemplo, G puede aparearse con U mediante un enlace menos fuerte que el enlace G=C. En otros casos, en el enlace de hidrógeno participa el 2'-OH de la ribosa. También pueden existir bases enfrentadas pero no complementarias, así como una o más bases a lo largo de la cadena separadas mediante un bucle para facilitar el acoplamiento de otras bases.

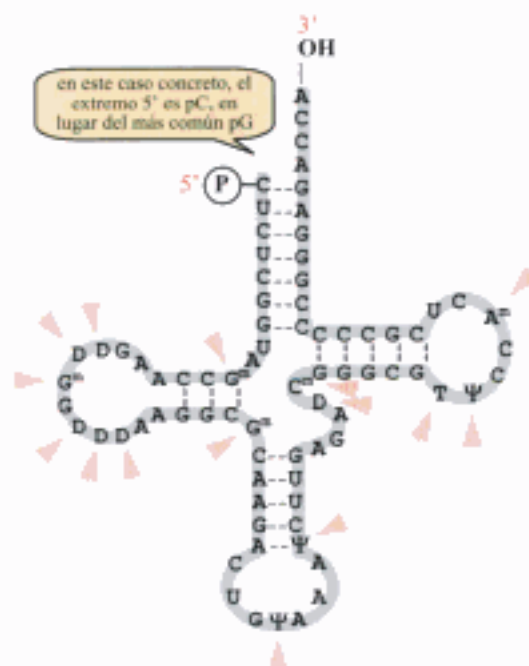
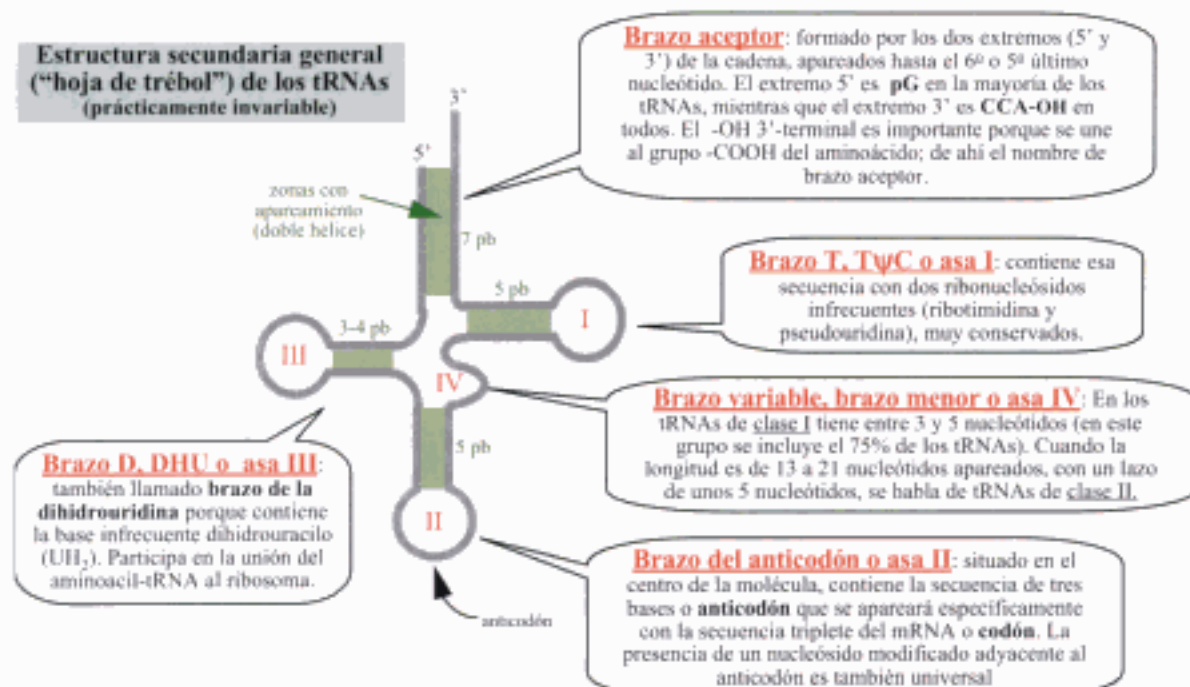


d) Estructura secundaria en trébol

Al comparar la estructura de distintos tRNAs se observa una notable semejanza. El plegamiento más probable es la estructura secundaria denominada *en trébol* (por su analogía con tres hojas y un tallo, o para otros con cuatro hojas), donde el apareamiento entre bases es máximo. En ella se distinguen:

- Tres **brazos o asas principales**; una central, II, y dos periféricas, I y III (cada una con tallo y bucle)
- Un **brazo o asa menor (IV)**, de dimensión variable entre los diferentes tRNAs
- Un brazo abierto o **brazo aceptor**

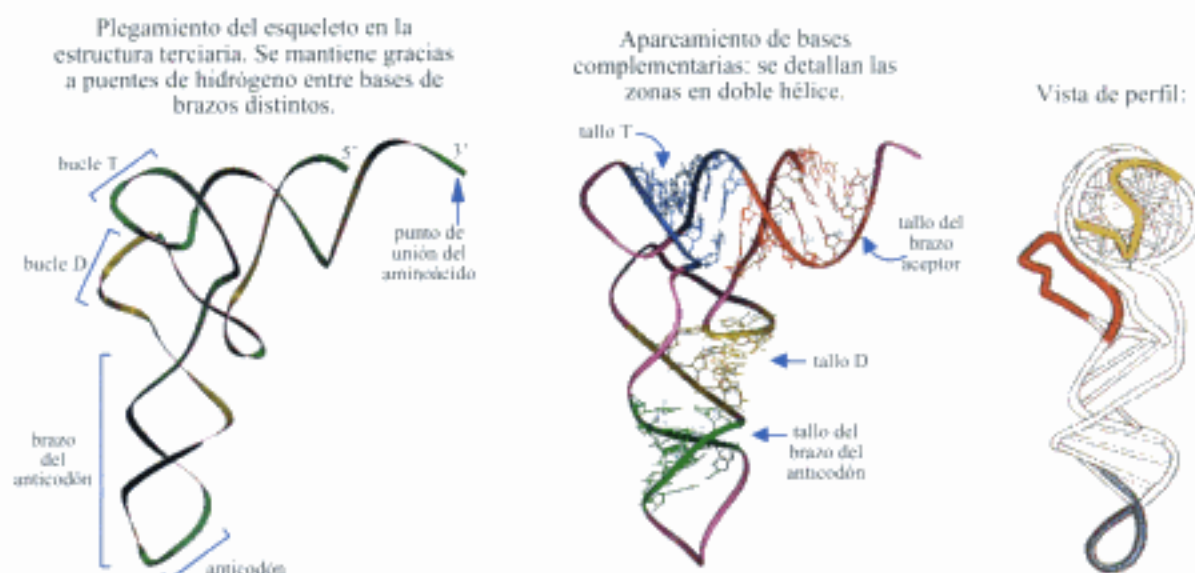
Estructura secundaria general ("hoja de trébol") de los tRNAs (prácticamente invariable)



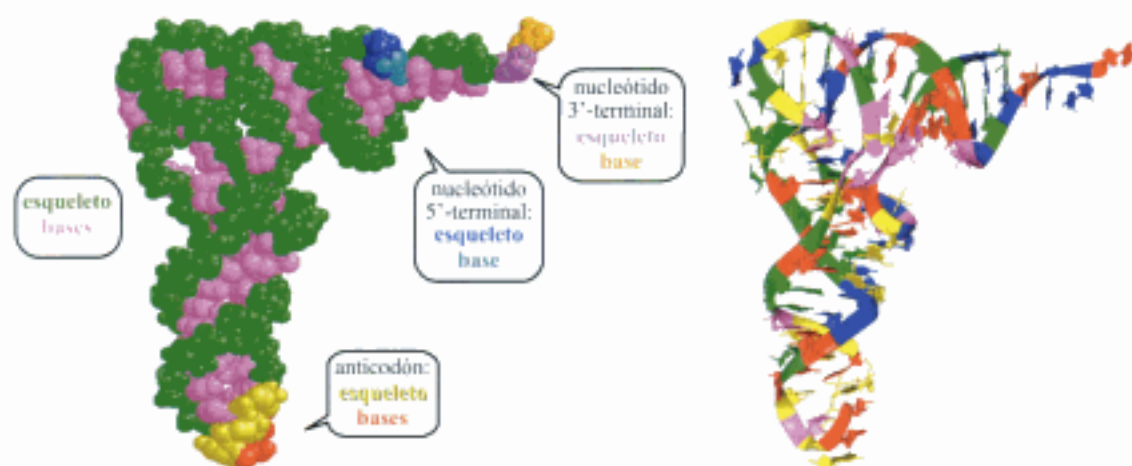
Ejemplo:
Detalles de la estructura primaria y secundaria del tRNA^{Tyr} de levadura

e) Estructura terciaria en "L"

Mediante análisis por difracción de rayos X de cristales de tRNAs se puede observar una estructura terciaria compacta, en forma de L. Ésta se origina por un retorcimiento en el espacio de la estructura secundaria en trébol, impuesto por las dobles hélices locales y por puentes de hidrógeno que acercan los brazos D y T. Estos puentes son de tipos infrecuentes, incluyendo en algunos casos tripletes de bases (pág. 73). La estructura en trébol es, por tanto, sólo una representación bidimensional simplificada de la verdadera estructura tridimensional.



Diversas representaciones de la molécula completa:

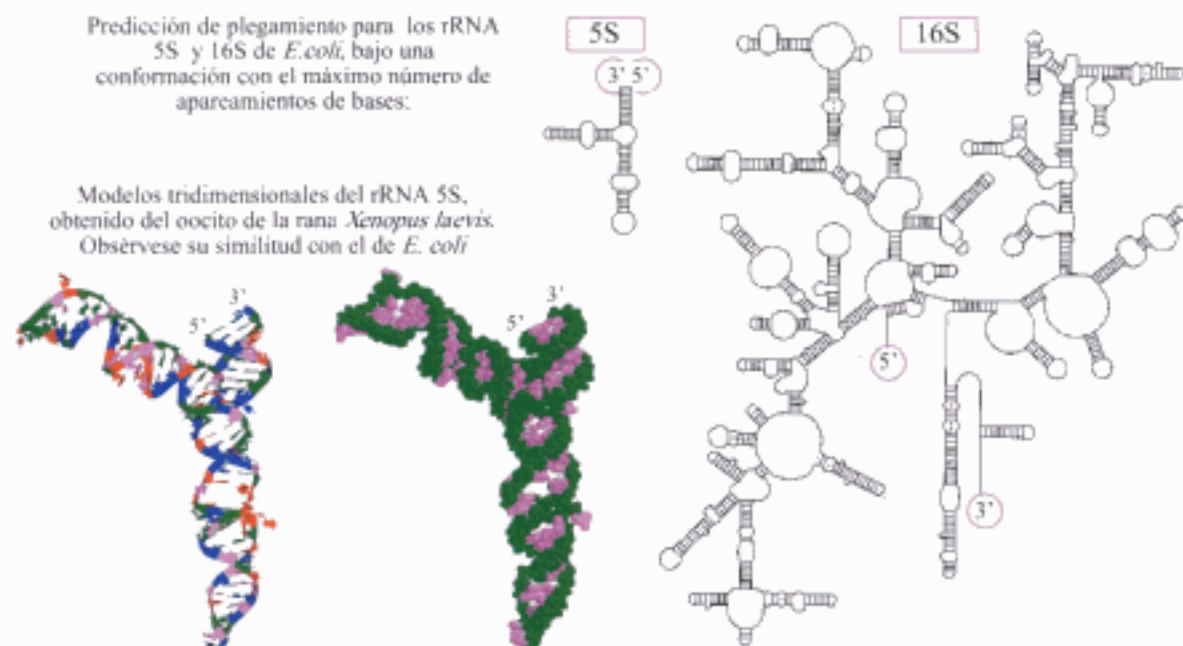


El conocimiento detallado de esta estructura permite comprender una peculiaridad de los tRNAs: todos son distintos (por tener distinta secuencia e interactuar con un codón de mRNA y un aminoácido específicos), pero al tiempo son de algún modo similares (por adoptar una estructura terciaria común, unirse a ribosomas y ser reconocidos por los factores que intervienen en la traducción).

7.2.3.3 RNA ribosómico (rRNA)

Es el **soporte estructural y componente principal de los ribosomas**. Existen distintos tipos de rRNA en cada célula procariota y eucariota, con masas moleculares diversas, constituyendo la mayor parte del RNA total (75%). Al igual que el tRNA, presentan un alto contenido de nucleósidos infrecuentes (Ψ, T y aquéllos con bases metiladas, fundamentalmente).

Los RNAs ribosómicos adoptan una conformación compleja. Están formados por una sola hebra con regiones helicoidales cortas resultantes del apareamiento intracatenario de bases. Son moléculas con muchos repliegues sobre sí mismas, muy flexibles y deformables, que se comportan hidrodinámicamente como espirales al azar, especialmente a baja fuerza iónica o alta temperatura. Parece que su estructura constituye un armazón o molde sobre el que se ensamblan varias proteínas para dar lugar al ribosoma.



En **eucariotas**, la mayor parte del RNA presente en el núcleo son los *transcritos primarios*, moléculas obtenidas directamente de la transcripción del DNA y que se convertirán en RNA maduro tras sufrir el proceso de maduración postranscripcional. El tamaño del mayor transcrito primario precursor del rRNA varía entre 6 y 14 kb, según la especie (con un coeficiente de sedimentación de 35 a 47S). Poco después de ser sintetizado, ese pre-rRNA es escindido para dar tres rRNAs maduros, de 28S, 18S y 5,8S; éstos, junto al rRNA 5S (producto de un gen independiente y, por ello, de un pre-rRNA diferente) son los RNAs que constituyen la mayor parte de los ribosomas eucarióticos (pág. 80).

En **procariotas**, un único pre-rRNA da lugar, también por procesamiento postranscripcional, a rRNAs maduros de 16S, 23S y 5S, que forman el ribosoma procariótico (pág. 80). El mismo precursor produce además algunos tRNAs.

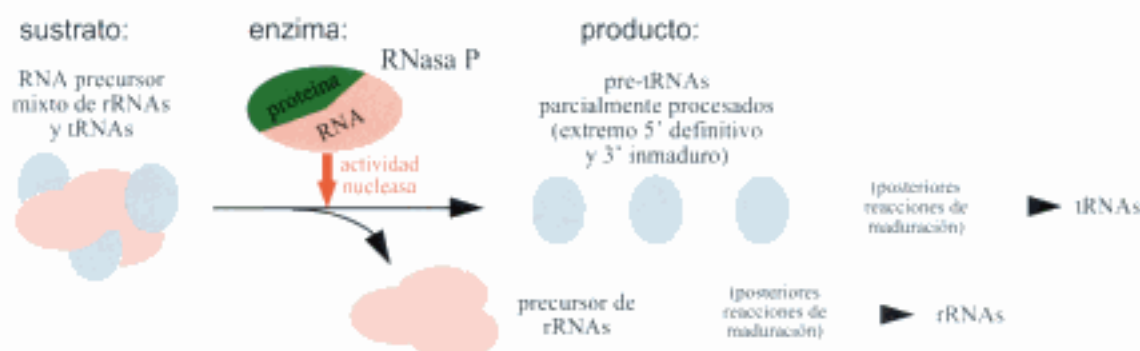
7.2.4 Ribozimas: función catalítica de algunos RNAs

Algunas moléculas de RNA son capaces de actuar como biocatalizadores; de ahí que el concepto clásico de enzima haya sido ampliado más allá de las proteínas. A estos RNAs con actividad catalítica, es decir, con capacidad de disminuir la energía de activación de determinadas reacciones bioquímicas, se les llama ribozimas, para distinguirlos de las enzimas proteicas. La capacidad catalítica estriba igualmente en la gran complejidad y diversidad de estructura tridimensional que pueden adoptar los RNAs, lo que les permite adaptarse al sustrato, así como en su reactividad gracias a los grupos funcionales de bases y esqueleto. De hecho, se han podido realizar estudios cinéticos con algunas ribozimas, verificando que cumplen la cinética de Michaelis-Menten, típica de enzimas proteicas. Se piensa que las ribozimas fueron los primeros biocatalizadores que aparecieron en la evolución, siendo desplazados posteriormente por las enzimas proteicas, más versátiles. Los RNAs serían así las moléculas primordiales de la vida, antes de que evolucionasen los DNAs como portadores de información genética y las proteínas como catalizadores.

Se conocen ribozimas que actúan en diversos tipos de reacciones de interés en Biología Molecular e Ingeniería Genética:

7.2.4.1 Acción catalítica en la maduración de RNA

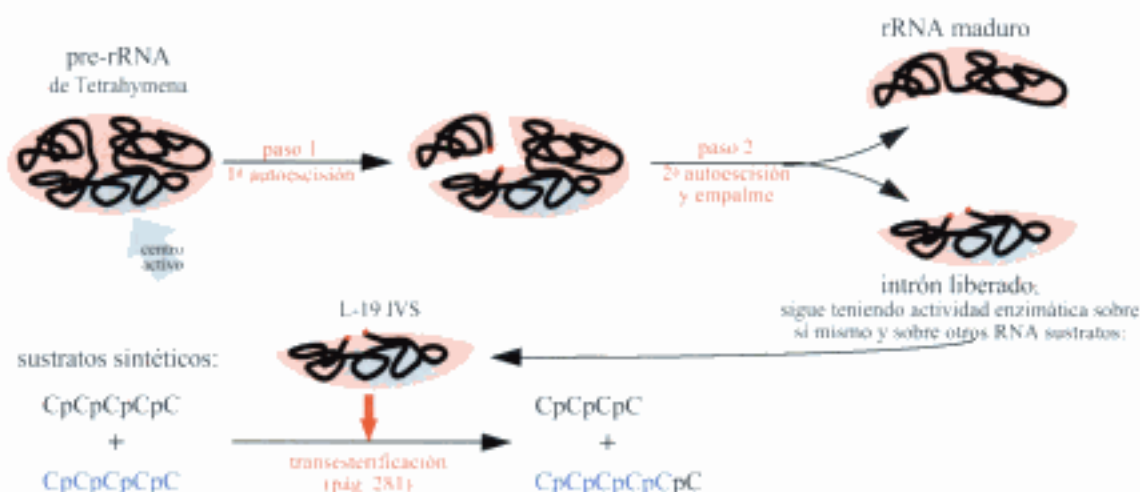
El primer ejemplo conocido de molécula de RNA con actividad catalítica fue la **RNasa P** (se ha estudiado en procariotas, pero parece existir una actividad similar en eucariotas). Esta enzima está formada por moléculas de proteína y de RNA, pero se ha comprobado que la actividad catalítica reside en el componente RNA. Reconoce como sustrato al precursor mixto de rRNAs y tRNAs en procariotas y cataliza una de las reacciones de su maduración postranscripcional (para dar los rRNAs y tRNAs funcionales), en concreto la generación de los extremos 5' de los tRNAs (pág.283).



7.2.4.2 Acción autocatalítica en la eliminación de intrones

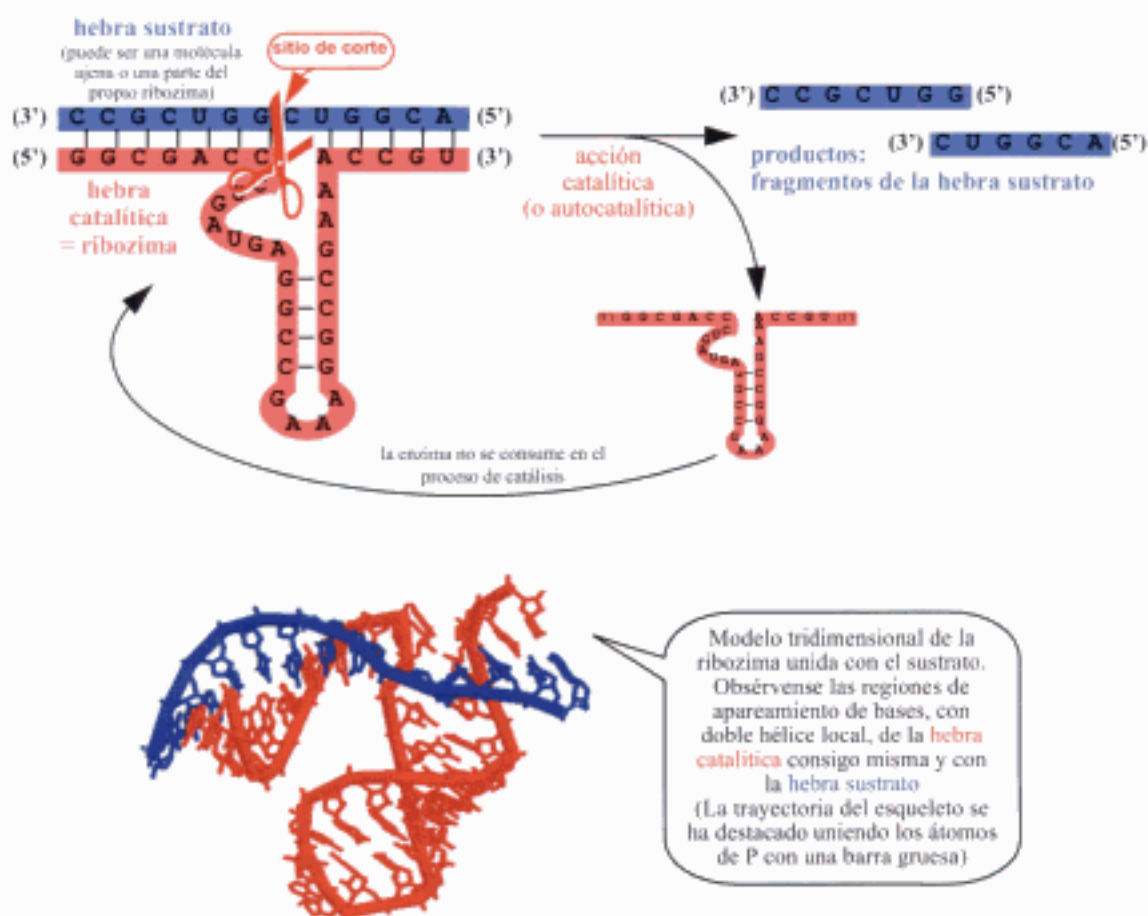
La siguiente "actividad ribozimática" encontrada fue la del **precursor de rRNA** en el protozoo *Tetrahymena* (eucariota unicelular). A diferencia del caso anterior, este ribozima actúa de forma autocatalítica, es decir, el sustrato es su propia molécula. Mediante un mecanismo similar actúan como ribozimas otros pre-RNAs procariotas y eucariotas, aquéllos que poseen **intrones de tipo I y II**. Los intrones (pág. 245) son secuencias del RNA precursor que no se traducen a proteína o no forman parte del RNA maduro, pues se eliminan durante la maduración; el pre-RNA comprende, pues, el RNA maduro y los intrones. Los tipos de intrón y los procesos de maduración del RNA precursor por corte y empalme se estudian en el procesamiento postranscripcional del RNA (pág. 280).

En este caso no se puede estrictamente hablar de catálisis, pues uno de los requisitos es que el catalizador no se consuma en el proceso. Sin embargo, el producto final de una serie de reacciones de autohidrólisis del intrón de *Tetrahymena* es una molécula (derivada del propio intrón y denominada L-19 IVS) que es capaz de catalizar reacciones de hidrólisis y esterificación de oligonucleótidos externos a ella. Por tanto, si bien no lo son el pre-RNA o el intrón originales, L-19 IVS sí es un verdadero biocatalizador.



7.2.4.3 Acción catalítica y autocatalítica en viroides: ribozimas de posible aplicación terapéutica

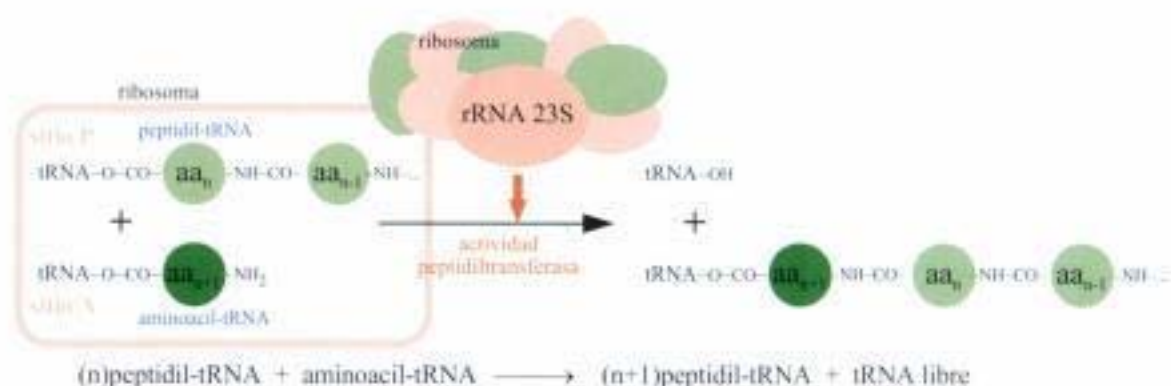
Los viroides son moléculas circulares de RNA monohebra con capacidad para infectar plantas. En algunos de ellos se han encontrado también ribozimas capaces de actuar, de forma análoga al caso anterior, sobre su propia molécula o sobre secuencias de RNA de moléculas diferentes. Estas ribozimas contienen dos componentes esenciales: por un lado, las secuencias de reconocimiento del RNA sustrato (por apareamiento de bases complementarias) y, por otro, su centro catalítico, que adopta una estructura tridimensional peculiar, denominada "cabeza de martillo" (*hammerhead*), y produce la escisión del RNA sustrato, inactivándolo.



Se han ensayado análogos sintéticos de alguno de estos ribozimas (obtenidos por ingeniería genética) para la destrucción selectiva de moléculas de RNA con secuencias características. Este es un planteamiento de terapia génica, dentro de la estrategia de "inhibición dirigida de la expresión génica". Se trata de preparar ribozimas artificiales con secuencias de reconocimiento que les permitan unirse "antisentido" a un mRNA, de forma que el centro catalítico lo escinda, suprimiendo la síntesis de la proteína correspondiente. (El concepto de cadena *antisentido* se introducirá en el tema de transcripción, pág. 248.) Otro tipo de planteamiento de terapia, dentro de la estrategia de "corrección dirigida de mutaciones", consiste en el corte mediante ribozimas de un transcrito precursor de RNA cuya secuencia estuviese mutada (por proceder de un gen mutado), de forma que la escisión ocurra alrededor de la mutación y resulte un transcrito corregido.

7.2.4.4 Acción catalítica transpeptidasa en la síntesis de proteínas

El cuarto ejemplo lo constituye la ribozima que ejerce la actividad transpeptidasa o **peptidiltransferasa**, responsable de la formación del enlace peptídico en la biosíntesis proteica, entre los residuos aminoácidos situados en los sitios A y P del ribosoma (pág. 313). En este caso, la "acción ribozimática" radica en el rRNA 23S, uno de los componentes de la subunidad mayor (50S) del ribosoma procariótico, y posiblemente también en el rRNA homólogo de eucariotas (rRNA 28S de la subunidad 60S).



A nivel fisiológico, la RNasa P y la peptidiltransferasa ribosómica son posiblemente los que mejor reflejan el parecido de las ribozimas con las enzimas proteicas; a saber: actúan sobre un sustrato específico, aceleran la reacción varios órdenes de magnitud, no se destruyen durante la catálisis y al parecer se comportan de acuerdo con la cinética michaeliana.

7.3 LOS RIBOSOMAS

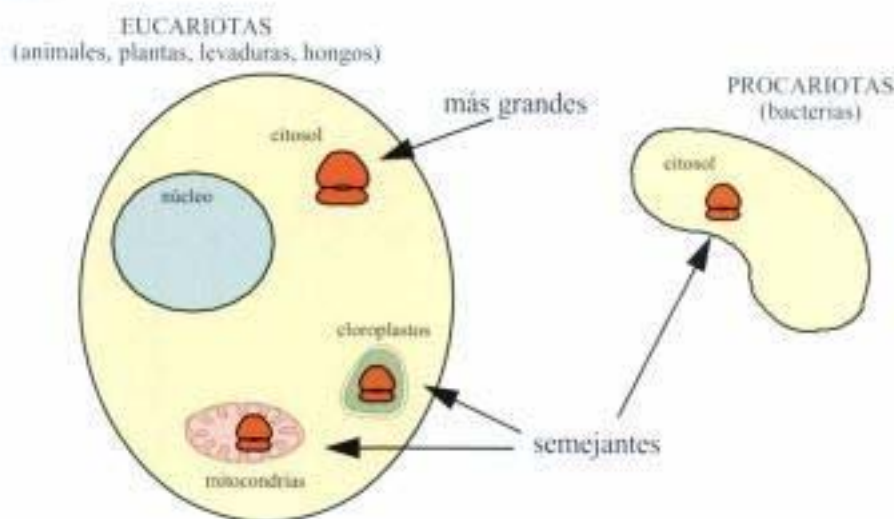
Los ribosomas son partículas subcelulares, u orgánulos, de importancia crucial para la transmisión de la información genética, al constituir el lugar físico donde se realiza la síntesis de proteínas. Se pueden considerar como estructuras químico-mecánicas pues, por un lado, se desplazan sobre el mRNA ordenando la interacción de sus codones con los correspondientes anticodones del tRNA y, por otro, proporcionan el entorno necesario para que los aminoácidos formen los enlaces peptídicos del polipéptido en crecimiento.

Estructuralmente, los ribosomas son **conjuntos plurimoleculares de proteínas y rRNAs** (35 y 65% en peso, respectivamente), con un tamaño de 20 a 23 nm. Es decir, son nucleoproteidos. El estudio de otros nucleoproteidos característicos, los virus, ya se aborda generalmente en una materia propia, la Virología.

7.3.1 Localización celular

En cada célula existen varios miles de ribosomas (por ej., más de 15.000 en un *E. coli*), que llegan a suponer un cuarto del peso seco de la célula. Se encuentran en el **citósol**, así como en la **matriz** mitocondrial y en el **estroma** de cloroplastos. Pueden estar en forma libre, como partículas discretas, pero a menudo se observan en forma de hilera, agrupados sobre un mRNA; en este caso se les llama **polisomas** o **polirribosomas**; su grado de agrupación o polimerización depende de la actividad metabólica celular.

Los ribosomas del citósol de eucariotas son más complejos en composición y mucho mayores en tamaño que los ribosomas bacterianos. Los ribosomas de mitocondrias y cloroplastos eucarióticos son semejantes a los bacterianos (citósólicos). Evolutivamente para todos los tipos se ha conservado una misma estructura secundaria, pero no la secuencia que la crea.



7.3.2 Características diferenciales de los ribosomas procarióticos y eucarióticos

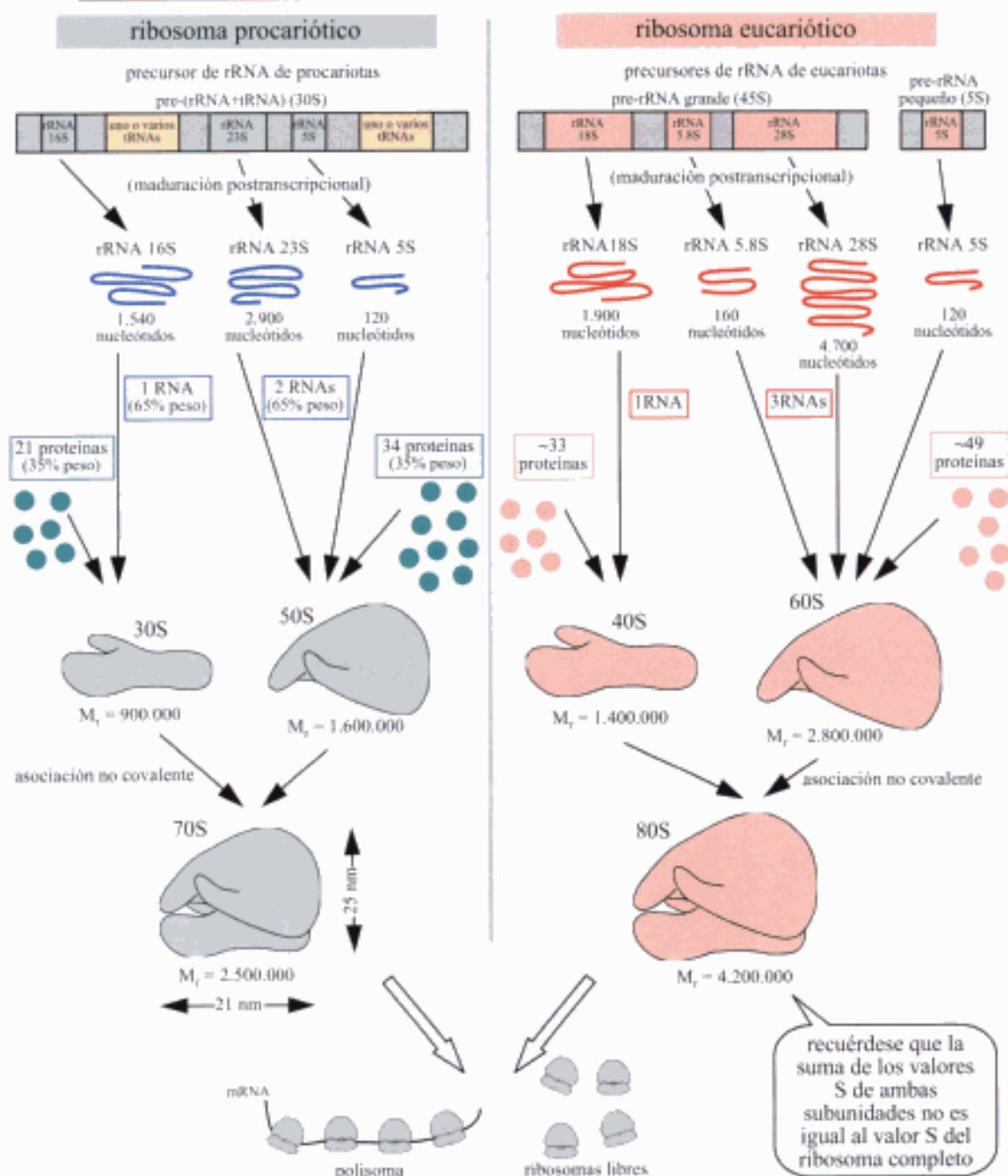
7.3.2.1 Origen

Los precursores del rRNA muestran una gran diversidad en su secuencia, que se manifiesta en la aparición (durante el proceso de maduración postranscripcional) de distintos tipos de rRNAs, caracterizados especialmente por su diferente coeficiente de sedimentación, *s*.

7.3.2.2 Composición

Los ribosomas están formados por **dos subunidades de tamaño diferente** y forma irregular, cada una constituida por uno o más rRNAs y varias proteínas de diversa masa molecular (entre 6 y 75 kDa). Tanto el ribosoma completo como sus subunidades y los rRNAs componentes se denominan según su coeficiente de sedimentación. Las proteínas, por su parte, se designan por números consecutivos precedidos por L (*large*) o S (*small*), según pertenezcan a la subunidad grande o pequeña, respectivamente.

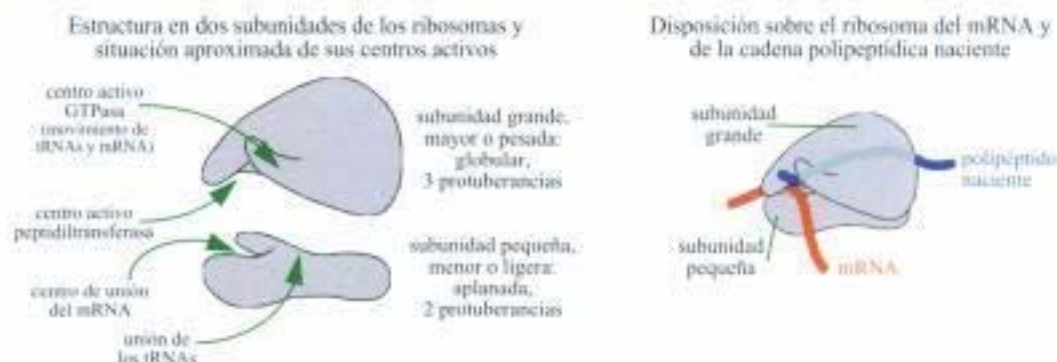
Diferencias en origen, composición y estructura de los ribosomas procarióticos y eucarióticos



7.3.2.3 Estructura

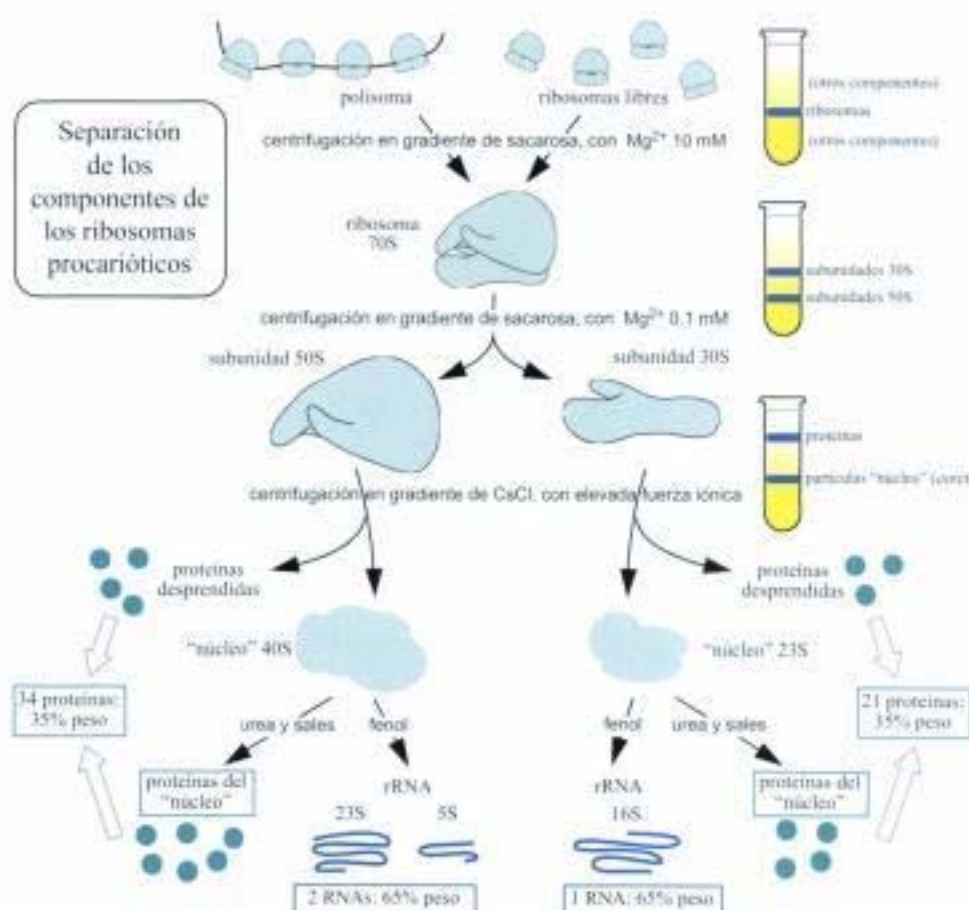
La estructura de los ribosomas, igual en todas las células, es muy característica; se ha dilucidado empleando difracción de rayos X, difracción de neutrones, microscopía electrónica en combinación con métodos inmunológicos, etc. Debido a su enorme tamaño (MDa) y la complejidad de su asociación interna, el estudio de la estructura, dinámica y función de los ribosomas constituyó un importante reto investigador.

Las dos subunidades **se asocian** por interacciones no covalentes, de una forma peculiar, como si la subunidad pequeña llenase el hueco de la grande, dejando una hendidura o túnel entre ambas por la que pasa el mRNA a medida que el ribosoma se desplaza durante el proceso de traducción y de la que emerge la cadena polipeptídica.



7.3.3 Separación de los ribosomas, subunidades y componentes

Tanto los ribosomas individuales como sus subunidades y sus componentes (rRNAs y diferentes proteínas) se pueden separar mediante centrifugación en gradientes de sacarosa o de cloruro de cesio, en presencia de concentraciones definidas de Mg^{2+} . Las subunidades aisladas no son activas. Por otro lado, si se mezclan los componentes (rRNAs y proteínas) a pH y fuerza iónica correctos, se asocian espontáneamente en unidades funcionales.



Tema 8

Condensación del DNA y cromosomas

8.1	Condensación del DNA en eucariotas	83
8.1.1	Proteínas componentes de la cromatina	84
8.1.1.1	Histonas	84
8.1.1.2	Proteínas no histonas	84
8.1.2	Niveles de condensación del DNA nuclear eucariótico	85
8.1.2.1	Disposición en nucleosomas y fibras de 10 nm	85
8.1.2.2	Formación de la fibra de 30 nm	87
8.1.2.3	Cromatina o cromosoma interfásico	87
8.1.2.4	Cromosoma metafásico	88
8.2	Estudio del cromosoma metafásico	88
8.2.1	Aspectos citogenéticos	88
8.2.1.1	Morfología de los cromosomas	88
8.2.1.2	Cariotipo y su análisis	89
8.2.1.3	Bandeado	90
8.2.2	Regiones con significado funcional	92
8.2.2.1	Centrómero	92
8.2.2.2	Telómeros	92
8.3	Dotación genética en eucariotas	93
8.3.1	Número de cromosomas (n): carácter haploide y diploide	93
8.3.2	Contenido de DNA: valor C	94
8.3.3	Locus y alelo	94
8.3.4	Genotipo, fenotipo y dominancia	95

8.1 CONDENSACIÓN DEL DNA EN EUCARIOTAS

Para completar la descripción de la estructura del DNA como material genético y como requisito previo al estudio del ciclo celular (tema 9), es conveniente analizar el complejo problema de la condensación del DNA nuclear, que ocurre a lo largo de la interfase de dicho ciclo. Este aspecto, junto al proceso enzimático de la replicación del DNA (tema 12) durante la fase S, es de la mayor importancia desde el punto de vista básico y de sus aplicaciones.

Como ya se ha indicado, la **condensación** del DNA comprende dos aspectos, teóricamente independientes pero funcionalmente relacionados y subordinados entre sí:

Condensación del DNA

Superenrollamiento propiamente dicho de la molécula de B-DNA (doble hélice dextrógira).

Es análogo al superenrollamiento negativo ya descrito para procariotas (pág. 66); se conoce peor en eucariotas por su mayor complejidad, debida en parte a la gran magnitud del genoma.

Empaquetamiento, plegamiento o compactación de la molécula de DNA debido a la asociación con proteínas, para dar lugar a nucleoproteínas.

Se estudia en este tema.

Este aspecto apenas existe en procariotas y orgánulos de eucariotas, donde no hay histonas y la condensación se basa sólo en el superenrollamiento de la doble hélice.

La **condensación** del DNA desde cromatina a cromosoma ocurre de forma gradual durante la fase G₂ del ciclo celular (tema 9), mientras que el proceso inverso, la **descondensación** del DNA, comienza después de la división celular (fase M) y continúa durante la fase G₁ hasta alcanzar de nuevo la condición difusa, que permite la replicación (fase S). No se conocen con precisión los aspectos moleculares de ambos procesos, aunque se asume una gran analogía y se considera, por tanto, que ambos transcurren por los mismos mecanismos, en sentido inverso.

8.1.1 Proteínas componentes de la cromatina

8.1.1.1 Histonas

Son las principales proteínas componentes del material genético de eucariotas (equivalentes en masa al DNA). Su función fundamental es **estabilizar la estructura del DNA**, contribuyendo a compactarla, para facilitar su empaquetamiento. Estructuralmente, las histonas constituyen una familia de proteínas semejantes, de tamaño relativamente pequeño y con elevado contenido (casi 1/3 del total) de aminoácidos básicos (pI de 9 a 10). Gracias a ello, muestran naturaleza policationica a pH fisiológico (numerosas cargas positivas en cada molécula) y se asocian fuertemente, mediante interacciones electrostáticas, con los grupos fosfato del esqueleto polianiónico del DNA.

	% Lisina	% Arginina	%(Lys+Arg)	Nº de aminoácidos	Masa molecular
H1	24,8-29,5	1,3-2,6	27,4-30,8	215-244	21,1-23,0 kDa
H2A	10,9	9,3	20,2	129	14,0 kDa
H2B	16,0	6,4	22,4	125	13,8 kDa
H3	9,6	13,3	22,9	135	15,3 kDa
H4	10,8	13,7	24,5	102	11,2 kDa

Existen **cinco tipos** principales de histonas bien caracterizadas, que se diferencian en su composición y, como se verá después, en su papel estructural en la organización del nucleosoma y cromosoma. La histona H1 se presenta con mayor variabilidad de secuencia y tamaño en distintas especies o tejidos, mientras que las otras cuatro están altamente conservadas. En algunos casos, H1 es sustituida por una variante, denominada H5, con especial afinidad por la cromatina, lo que refuerza el empaquetamiento y la inactividad transcripcional.

El grado de condensación del DNA se regula en parte mediante la acetilación y fosforilación de las histonas (pág. 326), afectando así a la accesibilidad del DNA para la transcripción; de este modo las histonas intervienen en el control de la expresión génica.

8.1.1.2 Proteínas no histonas

En la cromatina se pueden encontrar otras proteínas, menos abundantes, más variables entre tejidos y especies y en general peor caracterizadas:

a) Con función estructural

- **Proteínas básicas:**
 - **Protaminas:** Desempeñan la función equivalente a las histonas en tipos celulares concretos, particularmente el espermatozoide. Son proteínas pequeñas (50 aminoácidos, en algún caso) con elevada proporción de Arg, Ala y Ser, y estructura predominantemente α -helicoidal. Su menor tamaño permite la compactación del DNA en el espacio mínimo disponible en el espermatozoide.
- **Proteínas ácidas:**
 - **HMG:** Hasta un 5% de las proteínas nucleares corresponden al grupo llamado HMG (*high mobility group*), proteínas del grupo de alta movilidad electroforética, por su carácter ácido, relativamente pequeñas. Algunas se asocian al nucleosoma (al menos *in vitro*), mientras que otras interaccionan con el DNA espaciador (pág. 86). Este grupo incluye algunos factores de transcripción.
 - **Proteínas del esqueleto o matriz nuclear:** Esta estructura, que sirve de soporte o guía en el empaquetamiento de la cromatina (pág. 87), está formada por diversas proteínas, incluyendo algunas HMG y otras proteínas de carácter ácido.

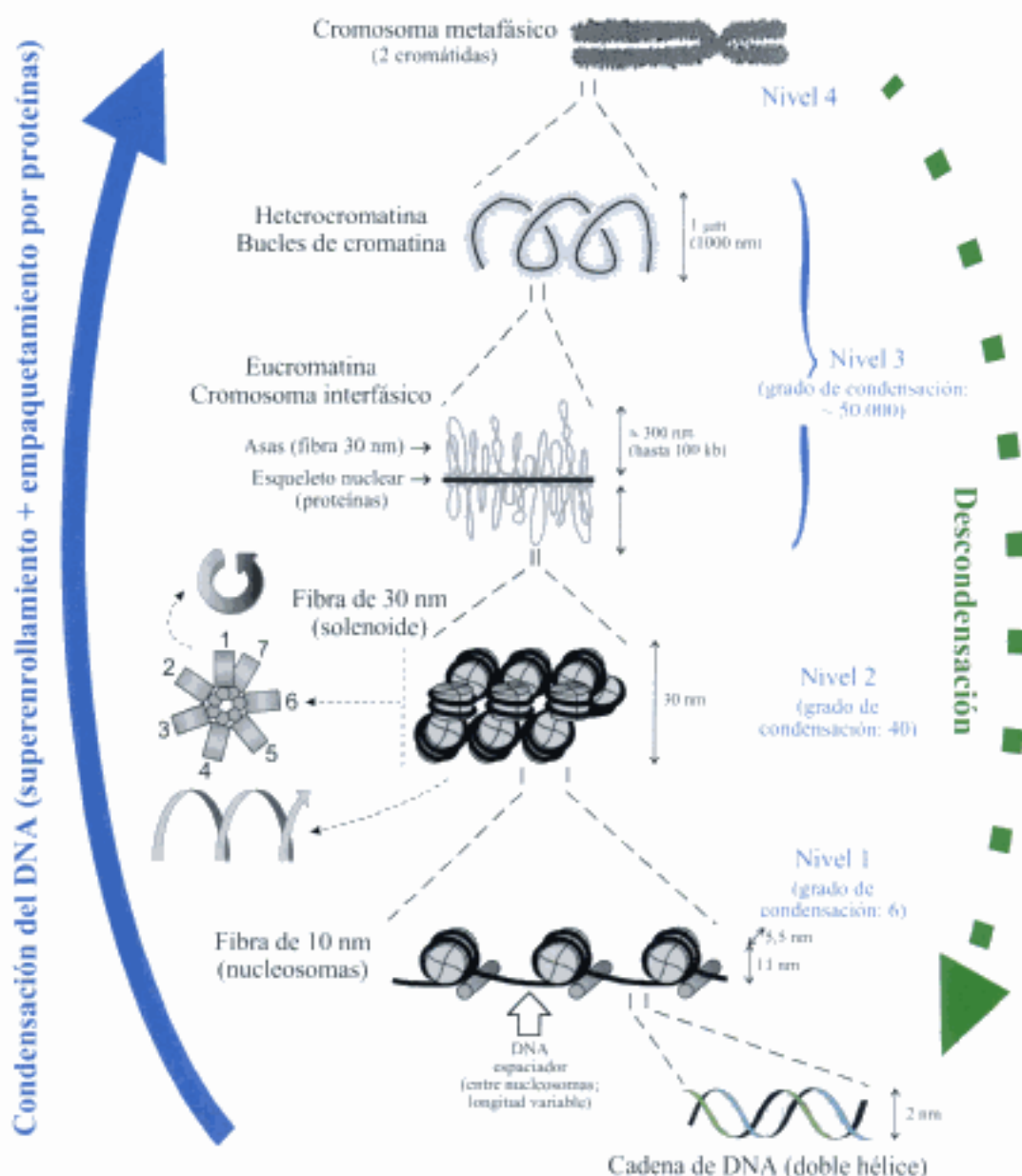
b) Con otras funciones

- Asociadas a la cromatina, aunque en mínima cantidad, se pueden encontrar numerosas proteínas, muy diferentes, que intervienen en la replicación y transcripción (temas 12 y 19) y en la propia regulación del grado de condensación de la cromatina: DNA- y RNA-polimerasas, sus proteínas auxiliares, factores de replicación, factores de transcripción, receptores hormonales, topoisomerasas, etc.

8.1.2 Niveles de condensación del DNA nuclear eucariótico

Se puede resaltar el grado de condensación alcanzado recordando que la longitud del DNA es varios órdenes de magnitud superior a las dimensiones de las estructuras celulares que lo albergan (pág. 66). Se denomina **grado de condensación o de empaquetamiento** al cociente entre la longitud del B-DNA en doble hebra y el tamaño del mismo una vez condensado; este parámetro se utilizará para comparar los distintos niveles de compactación.

La condensación del DNA puede estudiarse en cuatro niveles, aunque esta división es relativa, ya que todavía no se comprende con detalle cómo acontece este fenómeno a nivel molecular. Es importante indicar que la condensación y descondensación a través de esos niveles está estrechamente asociada a la progresión por las distintas fases del ciclo celular (págs. 100-101).

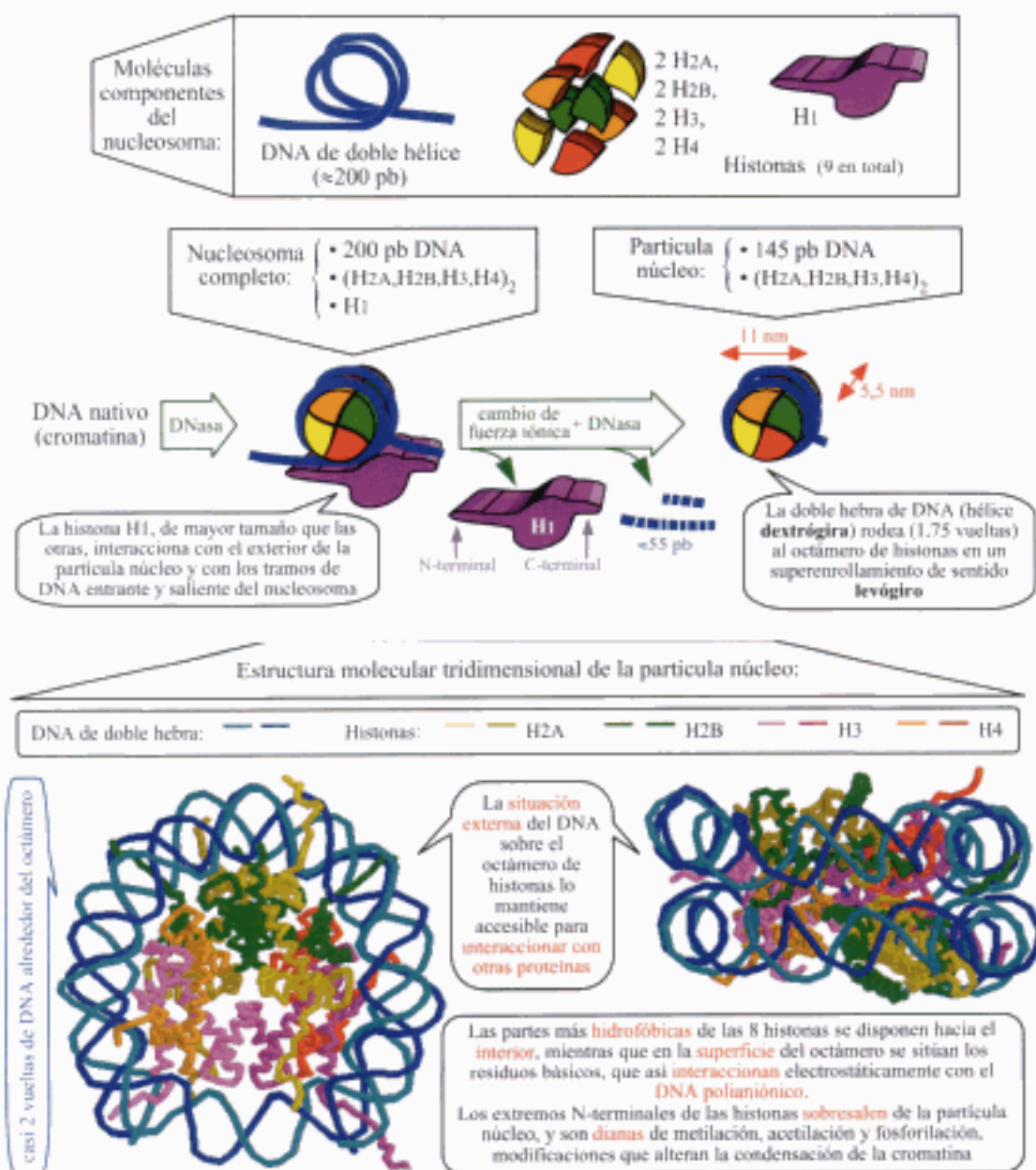


8.1.2.1 Disposición en nucleosomas y fibras de 10 nm

El nucleosoma es la unidad básica o elemental de la cromatina y los cromosomas. Constituye un patrón regular de interacción entre DNA e histonas, responsable de la organización estructural del material genético en todos los organismos eucariotas (en cualquier tipo celular y etapa del ciclo). Tanto el nucleosoma como sus componentes y las fibras en las que se integra han sido obtenidos y analizados por digestión con DNAsas, técnicas de entrecruzamiento (*crosslinking*), hibridación, difracción de rayos X, microscopía electrónica, difracción de neutrones, dicroísmo circular, ultracentrifugación analítica, etc. La propuesta de la existencia del nucleosoma surgió del estudio de la acción de DNAsas: sobre un DNA nativo, el tamaño de los fragmentos producidos es siempre múltiplo de 200 pb, mientras que si

actúan sobre un DNA previamente liberado de histonas, los fragmentos son de cualquier tamaño. Esto sugirió que la unión de las histonas protege regiones de DNA de unos 200 pb.

Cada **nucleosoma** está formado por 9 moléculas de histonas y un tramo de DNA. Bajo ciertas condiciones experimentales puede liberarse la histona H1, localizada externamente, lo que hace accesible a la DNasa una parte adicional del DNA. El nucleosoma queda así reducido a la llamada **partícula núcleo** (core), formada por un octámero de histonas rodeado por DNA. El arrollamiento levógiro del DNA sobre las histonas supone un superenrollamiento negativo de la cadena de DNA (en procariotas el superenrollamiento se consigue por simple retorcimiento del DNA circular, sin implicación de proteínas, y no existen nucleosomas).



Finalmente, otras proteínas no histonas también participan en la estructura: HMG-14 y HMG-17 se asocian, al menos *in vitro*, con cada nucleosoma, mientras que HMG-1 y HMG-2 interactúan con el DNA espaciador.

Una misma molécula de DNA envuelve sucesivamente a distintos nucleosomas y los conecta entre sí. La porción presente entre dos nucleosomas sucesivos se denomina **espaciador**, ligador o DNA internucleosómico (*linker*); su longitud puede variar entre 20 y más de 100 pb, dependiendo del organismo o incluso del tejido dentro de un mismo organismo. Esta estructura de nucleosomas espaciados a lo largo de la molécula de DNA se denomina **fibra básica** de cromatina o **fibra de 10 nm**, pues su grosor corresponde al diámetro del nucleosoma. El grado de empaquetamiento alcanzado es unas 6 veces superior al de la doble hélice del DNA extendida. La fibra de 10 nm también se llama "estructura en cuentas de rosario", por su aspecto característico al microscopio electrónico, generalmente sólo observado cuando se prepara *in vitro* a baja fuerza iónica (por ejemplo, 1 mM), mientras que dentro de la célula la cromatina se encuentra mayoritariamente en formas más condensadas, y su disposición en este estado relativamente extendido se limita, en todo caso, a pequeñas zonas.

8.1.2.2 Formación de la fibra de 30 nm

En las preparaciones de cromatina obtenidas *in vitro* a mayor fuerza iónica (por ejemplo, hasta 100 mM), la fibra de 10 nm se enrolla en sentido levógiro para dar lugar a una forma más condensada, denominada **solenóide** o **fibra de 30 nm**, que supone un grado de empaquetamiento 40 veces superior al de la doble hélice original. Esta fibra contiene algo más de 6 nucleosomas por cada vuelta de solenóide. Aunque se desconoce su estructura exacta *in vivo*, se cree que la histona H1 queda situada en el interior del solenóide, acercando los nucleosomas entre sí. De hecho, experimentalmente la formación de la fibra de 30 nm, a diferencia de la de 10 nm, requiere la presencia de H1. Esta es probablemente la forma fisiológica más descondensada de la cromatina, adoptada por la mayoría del DNA durante la interfase.

8.1.2.3 Cromatina o cromosoma interfásico

Este nivel de condensación supone la presencia de lazos y superenrollamientos de la estructura anterior. Los **bucles o asas radiales**, formados por unos 50 ó 100 kb de fibra de 30 nm, emergen de un armazón de proteínas no histonas denominado núcleo o andamiaje central, matriz nuclear o proteica, o **esqueleto nuclear**. Estos bucles sufren superenrollamiento, regulado por topoisomerasas asociadas al esqueleto. Parece que, además de contribuir al empaquetamiento del DNA, los bucles constituyen unidades transcripcionales, al haberse observado la presencia dentro de un solo bucle de un gen completo o bien de un grupo de genes regulados de forma común. Es típica la observación de asas radiales en la interfase de cromosomas politénicos gigantes de larvas de insectos y en cromosomas plumosos meióticos de oocitos de anfibios.

A partir de esta estructura se forman enrollamientos superiores, que hacen que la cromatina en la interfase presente en el núcleo un grado de condensación variable; se calcula que se multiplica el grado de condensación del DNA por 1.000 ó 2.000. Esta diversidad local en la densidad de condensación se observa en forma de la presencia simultánea de **euromatina** y **heterocromatina** en una célula en interfase.

	Heterocromatina	Euromatina (cromatina verdadera)
Aspecto al microscopio bajo tinción (definición original)	Se tiñe intensamente con el colorante May-Grünwald-Giemsa	Se tiñe débilmente y con aspecto difuso
Grado de condensación del DNA	Máxima condensación de la cromatina (próxima, aunque menor, a la del cromosoma metafásico)	Forma menos condensada
Contribución al total de la cromatina	Minoritaria; sólo algunas porciones del material genético se encuentran en esta forma: Heterocromatina constitutiva: regiones del genoma que se encuentran permanentemente en forma de heterocromatina (durante toda la interfase) y en todas las células Se descondensa el tiempo mínimo necesario para replicarse durante la fase S del ciclo celular. Incluye el <i>DNA satélite</i> (tema 10), concentrado especialmente en regiones teloméricas y centromérica de los cromosomas (pág. 92) Heterocromatina facultativa: se encuentra como heterocromatina sólo en algunas células de un mismo organismo o en algunos individuos de una especie; puede pasar a euromatina, en respuesta a determinadas señales Uno de los ejemplos es la inactivación de uno de los 2 cromosomas X en la mujer, puesto que se puede inactivar tanto el X materno como el paterno. Uno permanece condensado como heterocromatina, mientras el otro está en forma de euromatina Al final de la fase G ₂ , previa a la división por mitosis o meiosis, toda la cromatina de la célula adopta la forma de heterocromatina (antes de condensarse aún más para dar el cromosoma)	Mayoritaria durante la interfase
Accesibilidad de la molécula de DNA para su interacción con proteínas (DNA polimerasas, RNA polimerasas, factores de transcripción...; pág. 250)	No es accesible, debido a su elevada condensación. Se habla de cromatina <i>transcripcionalmente inactiva</i> . Se replica al final de la fase S. Los genes que contiene no se expresan	Es accesible: cromatina <i>transcripcionalmente activa</i> . Se replica al principio de la fase S. Sus genes se expresan

8.1.2.4 Cromosoma metafásico

Aunque al final de la profase cada cromosoma ya está condensado como heterocromatina, durante la metafase se condensa aún más, adquiriendo el aspecto típico de los **cromosomas metafásicos**, corpúsculos observables al microscopio óptico que muestran, en general, aspecto de bastoncillos con dos **cromátidas** más o menos separadas entre sí.

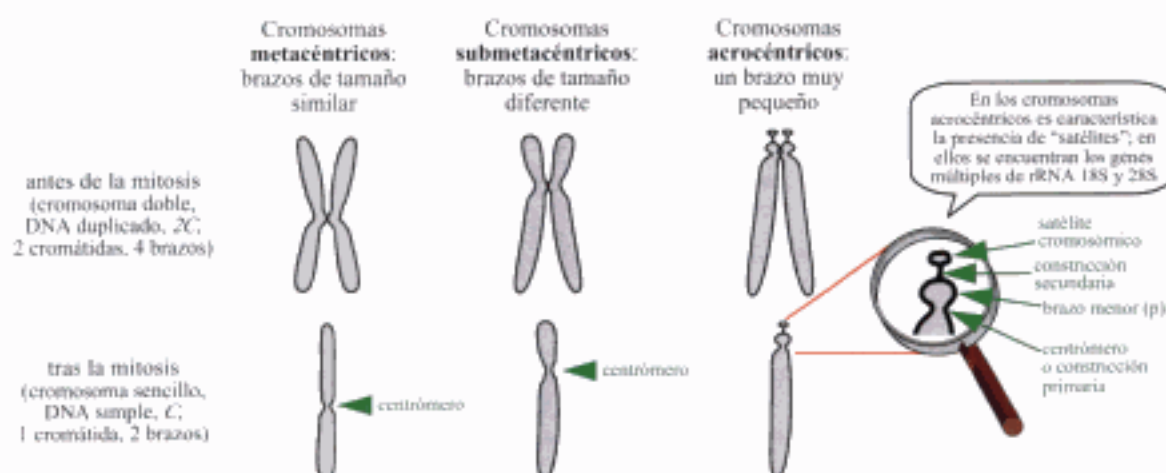
Las fibras de heterocromatina se arrollan en una primera espiralización para formar una **espiral menor** de paso muy pequeño, que se arrolla a su vez sobre sí misma (en unas 10-30 vueltas) para dar la **espiral somática**. Con esa doble espiralización se acorta la cromátida a una vigésima parte, sin aumentar su diámetro. Esto la hace más visible al microscopio óptico. Algunas regiones de la cromátida se arrollan sólo en espiral menor, como, por ejemplo, las del centrómero.

8.2 ESTUDIO DEL CROMOSOMA METAFÁSICO

8.2.1 Aspectos citogenéticos

8.2.1.1 Morfología de los cromosomas

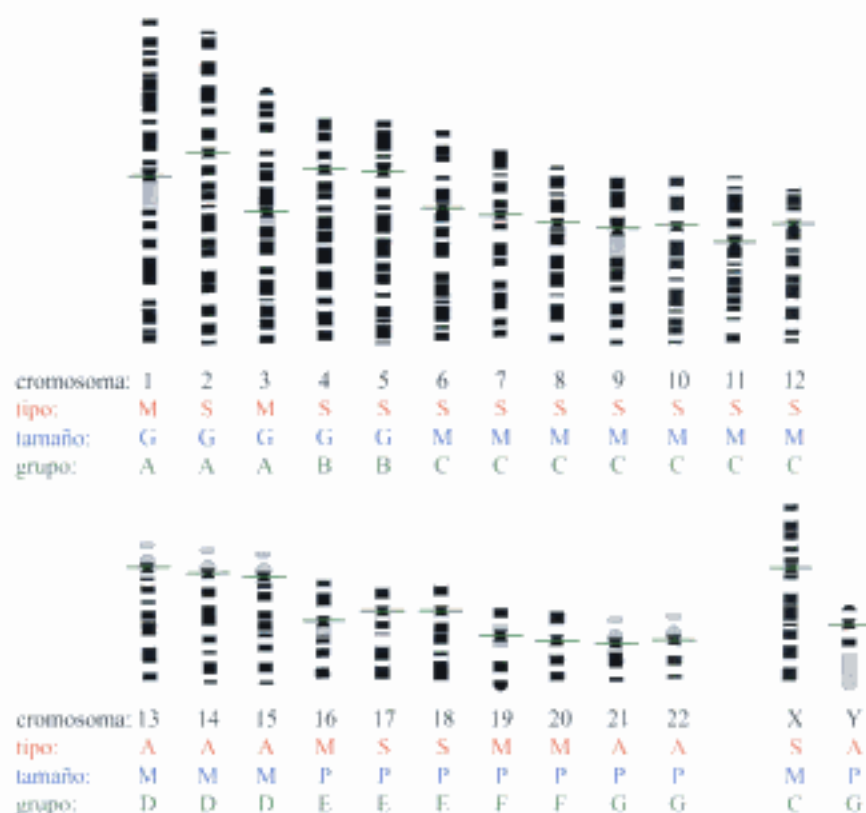
Se llama **centrómero**, **constricción primaria** o **central** a la región más estrecha del cromosoma, por la que permanecen unidas las dos cromátidas hermanas. El centrómero delimita los **brazos cromosómicos** (cuatro antes de la mitosis, dos tras ella). Los brazos cortos se designan con la letra **p** (de *petit*) y los largos con **q** (de *queue*). Cada brazo se nombra, pues, por el número del cromosoma al que pertenece, seguido de la letra **p** o **q**. Por ejemplo, **6p** es el brazo corto del cromosoma 6, **8q** el brazo largo del cromosoma 8, etc. Los cromosomas se clasifican atendiendo a la posición del centrómero y, por tanto, según el tamaño relativo de los brazos:



En los brazos cortos de los cromosomas acrocéntricos (excepto el Y) existen, además, otros estrechamientos de la cromátida, denominados **constricciones secundarias**, zonas en las que la espiral somática presenta un diámetro más reducido. Delimitan una sección terminal en el cromosoma, a modo de una pequeña esfera, a la que se llama **satélite cromosómico**.

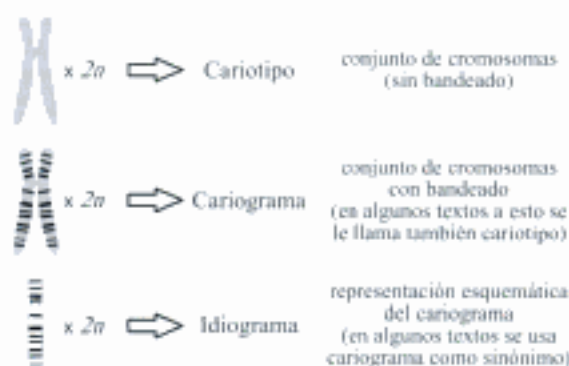
Clasificación de los cromosomas humanos, según:

- tipo (Metacéntrico, Submetacéntrico o Acrocéntrico)
- tamaño (Grande, Mediano, Pequeño)
- grupo de clasificación (A,B,C,D,E,F,G)



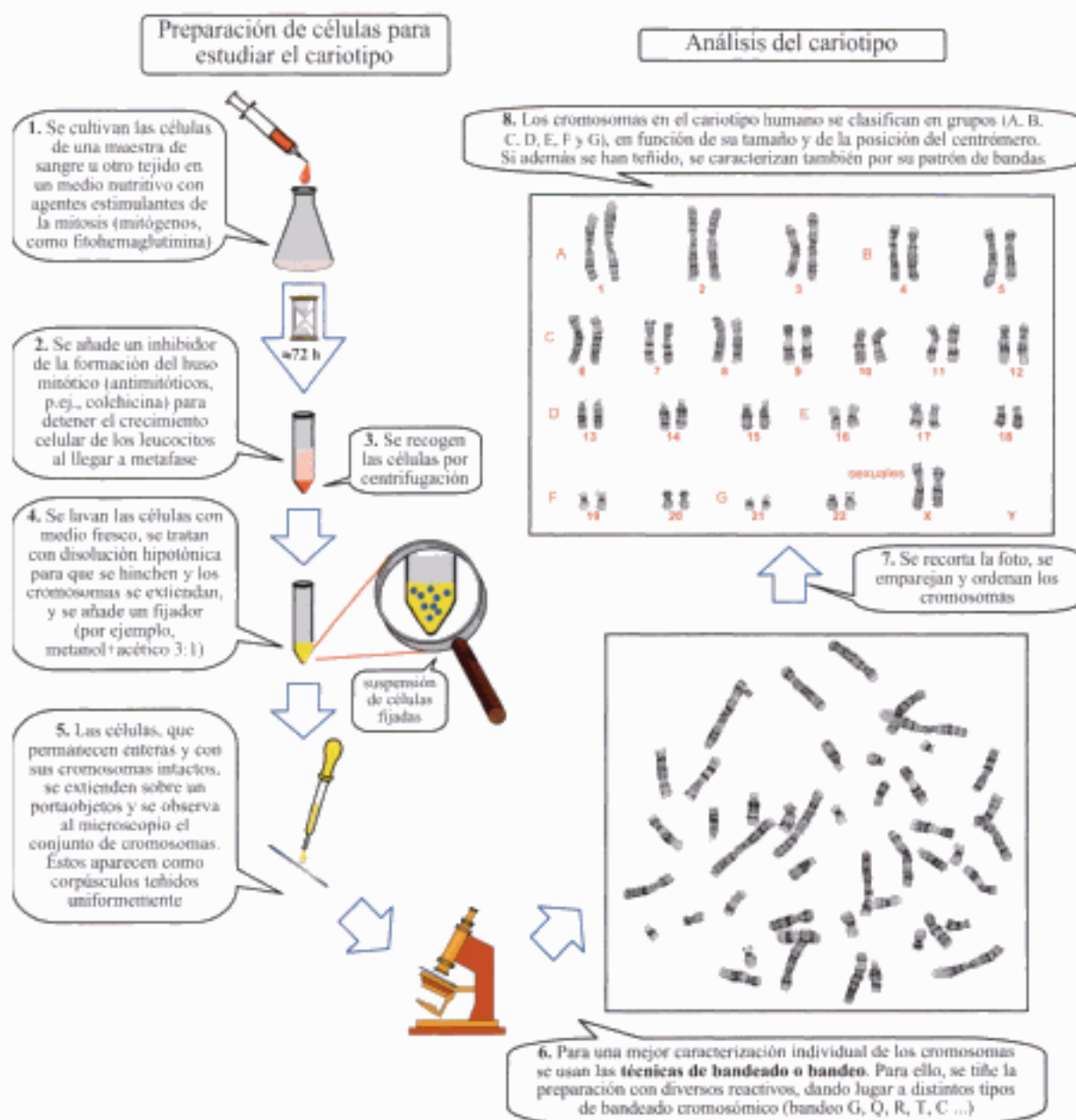
8.2.1.2 Cariotipo y su análisis

Cariotipo =
constitución cromosómica o complemento cromosómico de un individuo.
Es un rasgo característico de cada especie; todos los organismos de una especie tienen el mismo cariotipo. Sin embargo, especies muy similares pueden tener cariotipos muy diferentes.



Son varios los métodos desarrollados para el análisis del cariotipo. Normalmente, este estudio se realiza sobre cromosomas en el estado de prometafase o metafase. Las preparaciones de cromosomas metafásicos se emplean además para analizar anomalías cromosómicas, determinar la presencia o ausencia de cromosomas extra.

A este fin, el método original y más sencillo de obtención de cromosomas parte de leucocitos, como células nucleadas del organismo más asequibles.

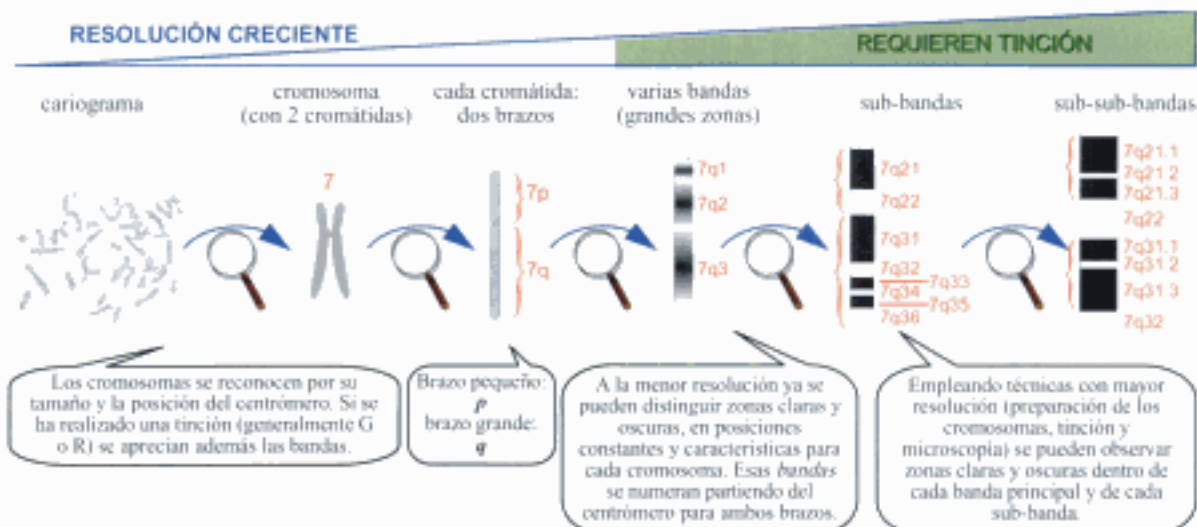


Actualmente, se parte ya no sólo de cultivos de leucocitos, con la desventaja de su vida corta, de 3 a 4 días, sino de cultivos a largo plazo de células de diversos tipos de tejidos (fibroblastos, médula ósea, líquido amniótico, biopsias, etc.). Las más usadas son las de mayor potencial de división o mayor facilidad de cultivo. Asimismo, existe una variedad de tratamientos de desnaturalización y/o degradación enzimática de la cromatina, y de tinciones con colorantes específicos para el DNA. Por último, se han desarrollado sistemas de cariotipificación que pueden ser informatizados de forma que seleccionen automáticamente las células a analizar, el tipo de tinción y el análisis de los cromosomas.

8.2.1.3 Bandedo

Aplicando distintas técnicas de tinción, utilizadas de forma rutinaria en laboratorios de citogenética, se puede observar en los cromosomas bandas pálidas y oscuras alternativamente, que definen grandes regiones cromosómicas llamadas **isocoros**. Este **bandeo o bandeo** de cromosomas no es consecuencia de agrupamientos fortuitos, sino que está relacionado con la organización estructural del genoma, reflejando variaciones tales como la composición de bases, grado de condensación cromosómica, conformación de la cromatina, secuencias repetitivas y no transcritas, etc.

Los patrones de bandas de cada cromosoma son prácticamente idénticos en células diferentes, y en casi todos los tejidos, pero pueden diferir entre individuos; por ejemplo, al menos 12 cromosomas muestran variaciones en la longitud de ciertos segmentos en distintas personas. Se pueden llegar a apreciar hasta 500 bandas en un cromosoma; en función de la resolución microscópica alcanzada, las **bandas** principales se subdividen sucesivamente en **subbandas** y **subsubbandas**. A todas ellas se les asigna un nombre, que incluye el cromosoma, el brazo y un número correlativo desde el centrómero hacia los extremos.



Para el bandeo de cromosomas se emplean diversos métodos de tinción del DNA con colorantes específicos:

a) Bando G

Es la técnica más utilizada. Se basa en una desnaturalización controlada de las proteínas cromosómicas (generalmente por digestión con tripsina), seguida de tinción con **Giemsa** (de donde viene el nombre de bandeo G) y observación al microscopio (el reactivo de Giemsa es una mezcla compleja de colorantes). Cada par de cromosomas muestra así patrones característicos de tinción, con bandas oscuras (G-positivas o bandas G) y bandas claras (G-negativas).

b) Bando Q

Los cromosomas se tñen con un compuesto fluorescente, como mostaza de **quinacrina** (hidrocloruro de quinacrina), 4',6-diamidino-2-fenil-indol (DAPI) o Hoechst 33258, que se intercalan en el DNA doble. Se requiere, por tanto, un examen mediante microscopía de fluorescencia. Aparece un patrón específico de bandas brillantes (bandas Q o Q-positivas), que se corresponden casi exactamente con las bandas G (G-positivas, G-oscuras).

Las bandas G y Q contienen DNA que, en general, es algo más rico en pares AT (55-60%), pues la quinacrina se inserta preferentemente entre estos pares. Corresponden a cromatina altamente condensada, con DNA de replicación tardía dentro de la fase S del ciclo celular, relativamente inactivo transcripcionalmente.

c) Bando R

Se basa en un tratamiento de los cromosomas con calor (para desnaturalizar el DNA rico en AT), antes de la tinción con Giemsa. Resultan así bandas oscuras (bandas R) que coinciden con las bandas G o Q claras (el nombre deriva de patrón *reverso*). También se consigue el mismo patrón de bandas empleando olivomicina, un colorante con afinidad por los pares GC.

Las bandas R contienen DNA rico en GC (50-60%), de baja condensación cromatinica, que se replica en etapas tempranas de la fase S y es relativamente activo transcripcionalmente.

d) Bando T

Identifica un subconjunto de bandas R especialmente concentradas en los telómeros, las de tinción más intensa, y se visualizan empleando un tratamiento térmico particularmente severo antes de tñir los cromosomas con Giemsa o una combinación de tñiciones y fluorocromos.

e) Métodos especiales

Otros métodos de cultivo cromosómico y de tinción se utilizan para situaciones particulares: **bando C**, para regiones heterocromáticas; **tinción NOR**, para la *región organizadora del nucleolo*, que contiene los genes de rRNA 18S y 28S; **bando de profase y prometafase**, o de **alta resolución**, sobre cromosomas detenidos en una etapa temprana de la mitosis, con un estado relativamente descondensado (su mayor longitud permite apreciar más detalles, llegándose a observar hasta 850 bandas en un solo cromosoma), etc.

Finalmente, la **citogenética molecular** emplea sondas de DNA clonado (preparadas por técnicas de DNA recombinante, pág. 176) para examinar cromosomas, de forma similar a los ensayos de hibridación molecular pero sobre preparaciones de cromosomas completos. Ésta es la técnica denominada **hibridación in situ con fluorescencia** (**FISH**, *fluorescence in situ hybridization*, pág. 172). Se dispone de sondas específicas para cromosomas individuales y para regiones cromosómicas, que permiten, por ejemplo, identificar reordenamientos cromosómicos particulares u obtener un diagnóstico rápido de la existencia de un número anormal de cromosomas.

8.2.2 Regiones con significado funcional

Los cromosomas eucarióticos muestran tres elementos esenciales para una función correcta del ciclo celular y la división, imprescindibles para la expresión, duplicación, y segregación de los cromosomas: el **centrómero** (lugar de unión del cromosoma a las fibras del huso acromático), los **telómeros** (regiones situadas en los extremos de los cromosomas) y los **orígenes de replicación** (puntos numerosos en cada cromosoma donde se inicia la copia de las dos hebras de DNA). Se tratan a continuación los aspectos morfológicos relacionados con los dos primeros; el análisis de los orígenes de replicación se pospone hasta que se estudie el proceso de replicación (pág. 150).

8.2.2.1 Centrómero

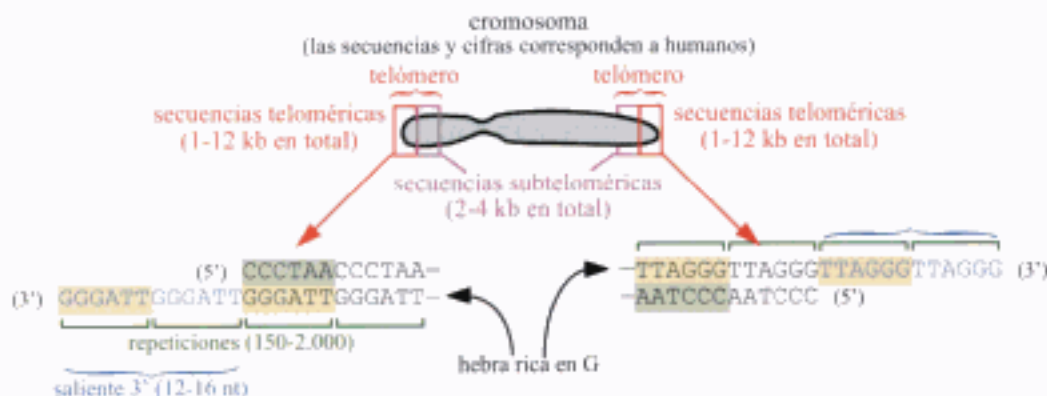
Se puede considerar el centro cinético del cromosoma, pues sobre él se sitúan las estructuras llamadas **cinetocoros**, a las cuales se unen las fibras que constituyen el huso acromático o huso mitótico. Estas fibras, filamentos contráctiles o microtúbulos, de naturaleza proteica, ejercen la tracción necesaria para separar las dos cromátidas durante la división celular. De esta forma se produce la segregación ordenada de los cromosomas y cada célula hija recibe una cromátida de cada cromosoma, es decir, idéntica dotación genética. La ausencia de centrómero (cromosoma **acéntrico**) impide que el cromosoma se una al huso mitótico y, por tanto, que se incluya en el núcleo de las células hijas.

Las secuencias esenciales para la función de los centrómeros son muy ricas en pares AT, tienen unos 170 pb en humanos y se repiten entre 2.000 y 30.000 veces en cada centrómero. Forman, pues, parte del llamado DNA repetitivo (pág. 113) y además suponen una parte importante de la fracción de heterocromatina constitutiva, aquella que sólo llega a descondensarse un tiempo mínimo en el ciclo celular, el preciso para su replicación.

8.2.2.2 Telómeros

Telómero es una palabra de etimología griega: *telos*, extremo, y *meros*, parte o región, que hace referencia a regiones situadas en los extremos de los cromosomas eucarióticos, constituidas por secuencias especializadas de DNA asociado a proteínas y con características estructurales y funcionales propias que las diferencian de otras regiones cromosómicas.

Los telómeros de la mayoría de eucariotas están formados por dos tipos de secuencias de DNA. Una de ellas, denominada **secuencia telomérica**, **repetición telomérica** o **repetición terminal**, constituye el extremo de la cadena de DNA del cromosoma, mediante la repetición en tándem de un oligonucleótido corto (variable según especies; en humanos es TTAGGG). Generalmente esta secuencia telomérica es más rica en G en una de las hebras, la que forma el extremo 3', que sobresale unos 12-16 nucleótidos sobre la hebra que aporta el extremo 5'. Ese extremo saliente es reconocido por **proteínas ligantes del telómero** (TBP, *telomere binding proteins*), que actúan a modo de "caperuza" protectora de dicho DNA terminal. Existen también unas secuencias más complejas adyacentes a las anteriores (es decir, más internas en el cromosoma), que se denominan **asociadas a telómeros** o **secuencias subteloméricas**.



Se han asociado a los telómeros distintas funciones:

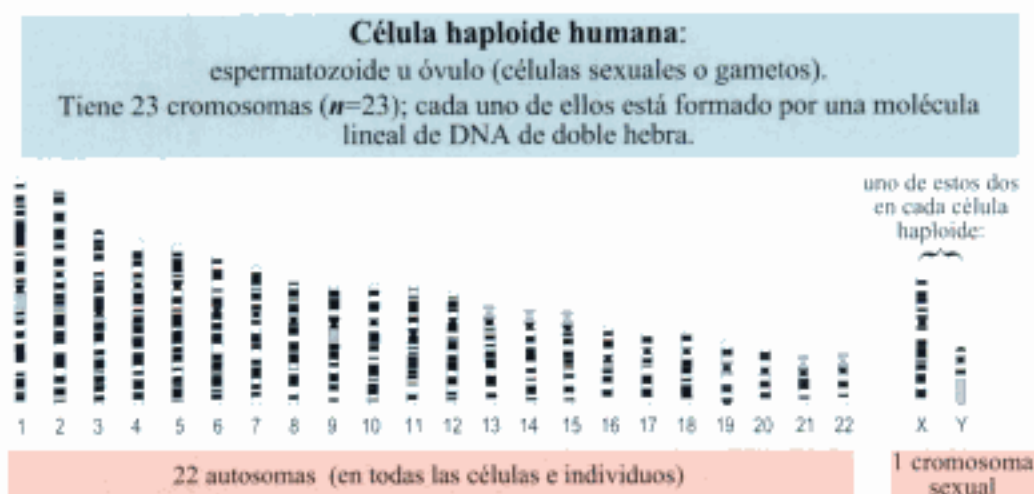
- Participar en la estabilización y mantenimiento de la integridad estructural del cromosoma: en ausencia de telómeros, el extremo del cromosoma tiende a unirse a otros y aumentan las posibilidades de que sufra recombinación y degradación. Por otra parte, los telómeros permiten a las enzimas encargadas de la reparación del DNA diferenciar entre el extremo natural del cromosoma y uno que resulte de la fragmentación accidental de la cadena de DNA; en este último caso se detiene el ciclo celular, evitando la replicación hasta que se haya reparado la lesión.
- La función más importante consiste en asegurar la replicación completa de los extremos del cromosoma. Este aspecto se estudiará una vez considerado con detalle el proceso de la replicación del DNA (pág. 157).
- Se especula también con la implicación del telómero en la arquitectura tridimensional del núcleo y/o del apareamiento cromosómico.

8.3 DOTACIÓN GENÉTICA EN EUCARIOTAS

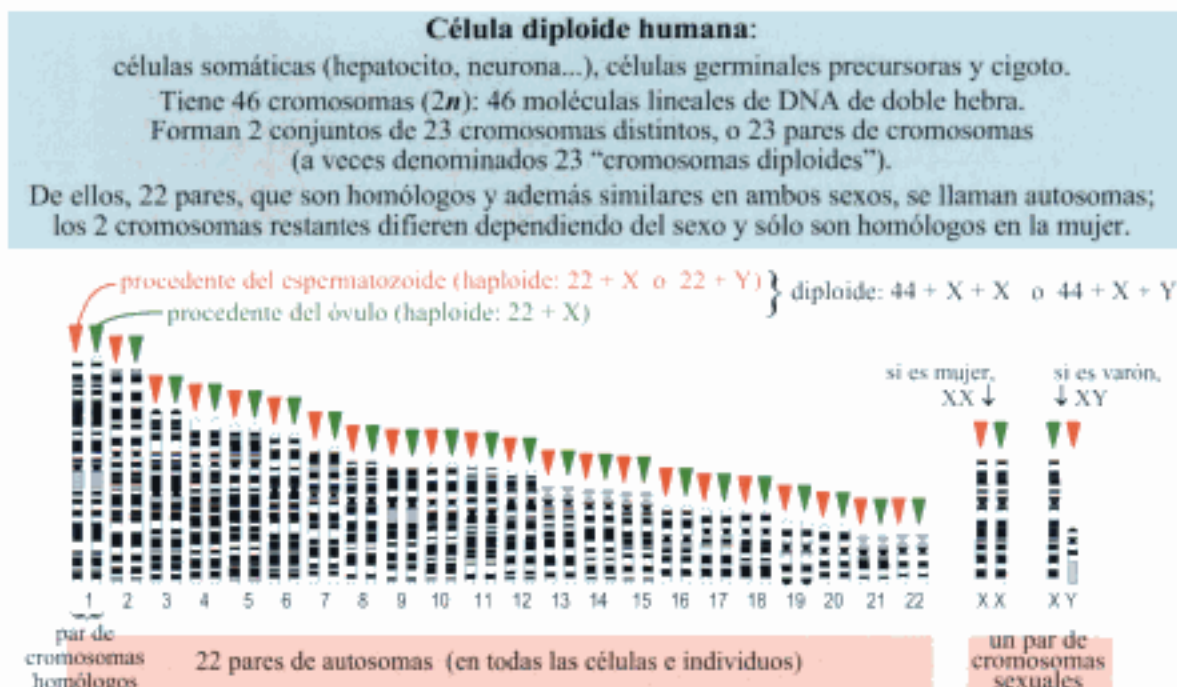
Es conveniente completar el conocimiento de la organización del material genético considerando el contenido total de este material presente en las células eucarióticas y su relación con el tipo de célula y el estado en el que se encuentra. Para ello, es preciso describir el significado de algunos términos y conceptos genéticos básicos, necesarios también para la comprensión del ciclo celular y de los dos tipos de división, mitosis y meiosis (tema 9).

8.3.1 Número de cromosomas (n): carácter haploide y diploide

El número de cromosomas nucleares es característico de cada especie, y varía mucho de unas a otras. Se habla de células **haploides** (y de organismos haploides, si están formados sólo por ellas) cuando cada cromosoma se presenta en forma individual, de modo que cada célula contiene un *juego* de cromosomas, todos distintos entre sí. El número de cromosomas se identifica con el número n .



Por el contrario, se denomina células **diploides** (y organismos diploides) a aquéllas en las que se observan parejas o *pares de cromosomas* morfológicamente iguales entre sí; los dos miembros de la pareja se denominan cromosomas **homólogos**. Por tanto, cada célula diploide tiene dos *juegos* de cromosomas, y este carácter diploide se identifica con el número $2n$ (n parejas de cromosomas). En cada par, un cromosoma se ha heredado del padre y otro de la madre.



cromosomas sexuales	mujer:	varón:
Homología:	 X X, homólogos El par nº 23 está formado por dos cromosomas homólogos, igual que los autosómicos (pares nº 1 a 22)	 X Y, muy distintos El "par" nº 23 está formado por dos cromosomas no homólogos Sólo una pequeña región del Y y el X son homólogos (región pseudoautosómica)
Gametos producidos:	Se dice que la mujer es <i>homogamética</i> : todos sus óvulos llevan un mismo tipo de cromosoma, X	Se dice que el varón es <i>heterogamético</i> : la mitad de sus espermatozoides son portadores de X, la otra mitad de Y
Procedencia de sus cromosomas:	un X de la madre, el otro X del padre	el X siempre de la madre, el Y siempre del padre
Transmisión de los cromosomas:	cromosoma X a las hijas y a los hijos, indistintamente (por fertilización con un espermatozoide X o Y, respectivamente)	el X sólo a las hijas, el Y sólo a los hijos (al fertilizar cualquier óvulo, que siempre será X)
En algunas especies, a diferencia de los humanos, no existe cromosoma Y, sino que los machos tienen un solo cromosoma sexual X, sin pareja, mientras las hembras tienen dos cromosomas X homólogos.		

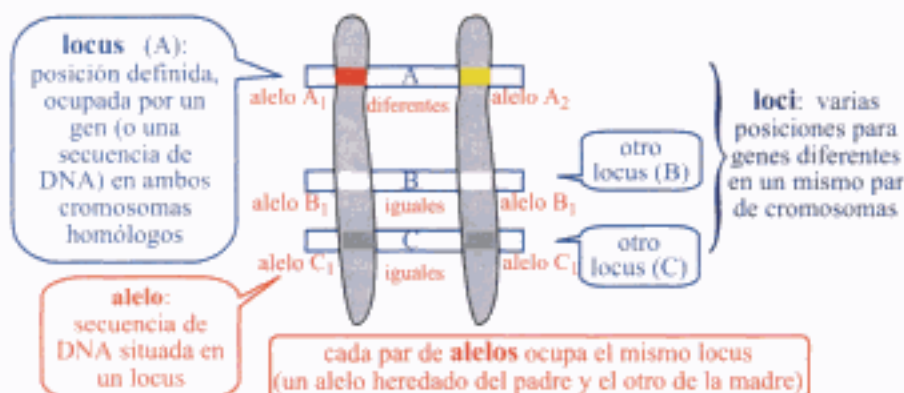
En el caso de células (u organismos) con más de $2n$ cromosomas ($4n$, $8n$, etc.) se habla de carácter **poliploide**. Hay células que son poliploides de forma natural, porque han formado copias adicionales de su juego de cromosomas inicial mediante una replicación sin división celular (endomitosis). Por ejemplo, las células del hígado en regeneración son tetraploides ($4n$), y los megacariocitos gigantes de médula ósea contienen usualmente $8n$, $16n$ o $32n$. Hay también situaciones patológicas caracterizadas por poliploidía: las más frecuentes son la triploidia ($3n$) y la tetraploidia ($4n$). Estas anomalías cromosómicas se estudian en el tema 29.

8.3.2 Contenido de DNA: valor C

Durante el ciclo celular, la formación de gametos y la fertilización tienen lugar variaciones en el número de cromosomas de la célula, indicadas mediante los múltiplos de n . En humanos y animales, estas variaciones transcurren entre n (haploide) y $2n$ (diploide). No debe confundirse esa variación con la del contenido en DNA de cada célula, aunque éste también varía durante los procesos citados. Para indicar este contenido se emplea una notación basada en denominar **C** a la cantidad de DNA (de doble hebra) correspondiente a un juego haploide de cromosomas (es decir, n) antes de su replicación. Se pueden encontrar células, tanto haploides como diploides, cuyo contenido C es diferente:

Célula haploide	n, C	gameto maduro (óvulo y espermatozoide)
	$n, 2C$	gametocito secundario (precursor de los gametos maduros)
Célula diploide	$2n, 2C$	célula somática cuando no está en división, antes de sufrir la replicación de su DNA
	$2n, 4C$	célula somática tras sufrir la replicación, a punto de dividirse por mitosis gametocito primario, antes de dividirse por meiosis dando 2 gametocitos secundarios

8.3.3 Locus y alelo

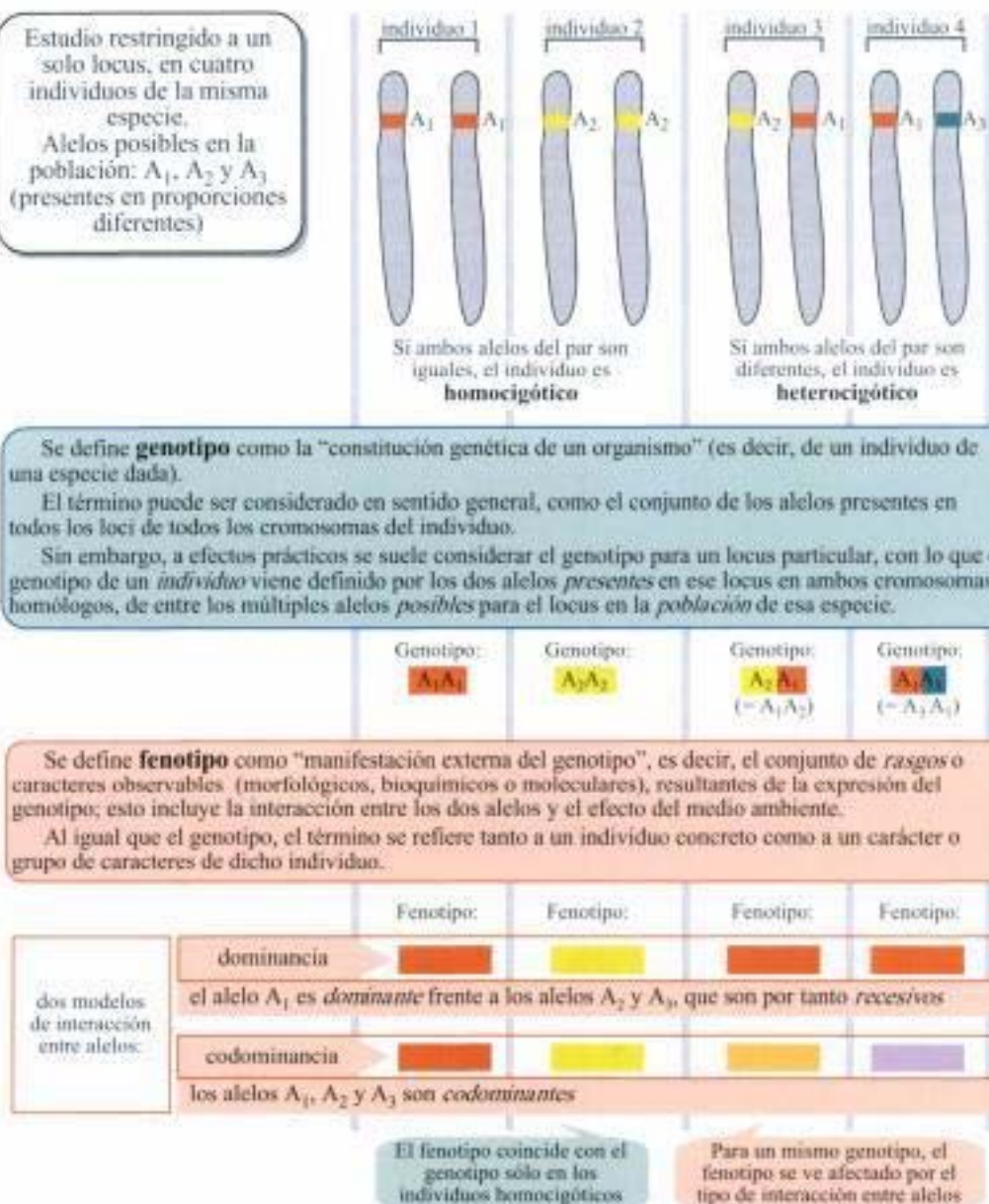


En cualquier célula de una misma especie, un gen ocupa siempre el mismo lugar en un cromosoma determinado, y la posición equivalente en su homólogo. (Por ahora, basta con que se entienda por gen una secuencia de DNA que codifica un producto funcional, bien proteína o bien RNA.) Para referirse de forma genérica a esa posición se utiliza el término **locus**. Por tanto, genes distintos ocuparán diferentes **loci** (plural de locus), en el mismo par de cromosomas o en otro. Estos términos también se utilizan para la posición de regiones de DNA que no sean genes.

A cada secuencia concreta de DNA (génica o no) situada en un locus se la llama **alelo**. Puesto que cada cromosoma se presenta como un par, en cada individuo hay dos alelos en el mismo locus, que pueden ser idénticos o diferentes.

En general, los dos alelos que presenta un individuo en un locus determinado son diferentes de los de otro individuo de la misma especie (diversidad genética o **polimorfismo**, pág. 365). Como consecuencia, en el conjunto de la población existen más de dos alelos diferentes para un locus. El estudio de trastornos clínicos ha permitido identificar numerosos alelos, aquéllos menos frecuentes en la población, que sirven hoy como marcadores para identificar dichas patologías.

8.3.4 Genotipo, fenotipo y dominancia



Relación entre el número de alelos existentes en la población para un solo locus y los genotipos diferentes a los que pueden dar lugar				Probabilidad de los genotipos en la población (asumiendo que todos los alelos son igualmente frecuentes)	
5 situaciones posibles en la población: N° de alelos existentes	N° de genotipos homocigóticos	N° de genotipos heterocigóticos	N° de genotipos posibles en total	Relación de homocigóticos a heterocigóticos	Frecuencia de heterocigosis
1: A_1	1: A_1A_1	0	1	1 : 0	0%
2: A_1, A_2	2: A_1A_1, A_2A_2	1: A_1A_2	3	2 : 2 ^(*)	50%
3: A_1, A_2, A_3	3: A_1A_1, A_2A_2, A_3A_3	3: A_1A_2, A_1A_3, A_2A_3	6	3 : 6 ^(*)	66%
4: A_1, A_2, A_3, A_4	4: $A_1A_1, A_2A_2, A_3A_3, A_4A_4$	6: $A_1A_2, A_1A_3, A_1A_4, A_2A_3, A_2A_4, A_3A_4$	10	4 : 12 ^(*)	75%
5: A_1, A_2, A_3, A_4, A_5	5: $A_1A_1, A_2A_2, A_3A_3, A_4A_4, A_5A_5$	10: $A_1A_2, A_1A_3, A_1A_4, A_1A_5, A_2A_3, A_2A_4, A_2A_5, A_3A_4, A_3A_5, A_4A_5$	15	5 : 20 ^(*)	80%
x	x	$x(x-1)/2$	$x(x+1)/2$	$x : x(x-1)$	$(x-1)/x \cdot 100\%$

^(*) Un mismo genotipo heterocigótico es el resultado de dos combinaciones de cada dos alelos (por ejemplo, $A_1A_2 = A_2A_1$, en la columna 3). Como consecuencia (columna 5), la probabilidad de existencia de cada uno de los genotipos heterocigóticos en la población es doble que la de cada genotipo homocigótico.

En las situaciones en las que existen dos o más alelos en la población se dice que hay **polimorfismo**, obviamente referido al locus considerado. Los polimorfismos del genoma humano son objeto de atención especial en el tema 26.

Tema 9

Ciclo celular

9.1	Introducción	97
9.2	La cromatina en la interfase	98
9.2.1	Variaciones en condensación y contenido en DNA de la cromatina	98
9.3	División de células somáticas por mitosis	99
9.4	División de células germinales primarias por meiosis	99
9.4.1	Gametogénesis	99
9.4.1.1	Espermatogénesis	102
9.4.1.2	Oogénesis	103
9.4.2	Reparto del material genético en la meiosis	104
9.4.2.1	Carácter reductor de la meiosis	104
9.4.2.2	Comparación entre meiosis y mitosis	104
9.4.2.3	Segregación de cromosomas y diversidad genética	105
9.4.2.4	Recombinación meiótica	106
9.5	Unión de los gametos haploides durante la fecundación	106
9.6	Resumen integrado	106

9.1 INTRODUCCIÓN

Es importante resaltar que los aspectos estructurales hasta ahora considerados para la molécula de DNA deben estudiarse bajo la situación fisiológica en la que realmente se encuentra el material genético en el núcleo eucariótico, es decir, durante las etapas de reposo y división que se alternan en la vida de la célula. Con frecuencia, estas dos cuestiones se relacionan poco entre sí, o se analizan bajo planteamientos sin una suficiente base molecular.

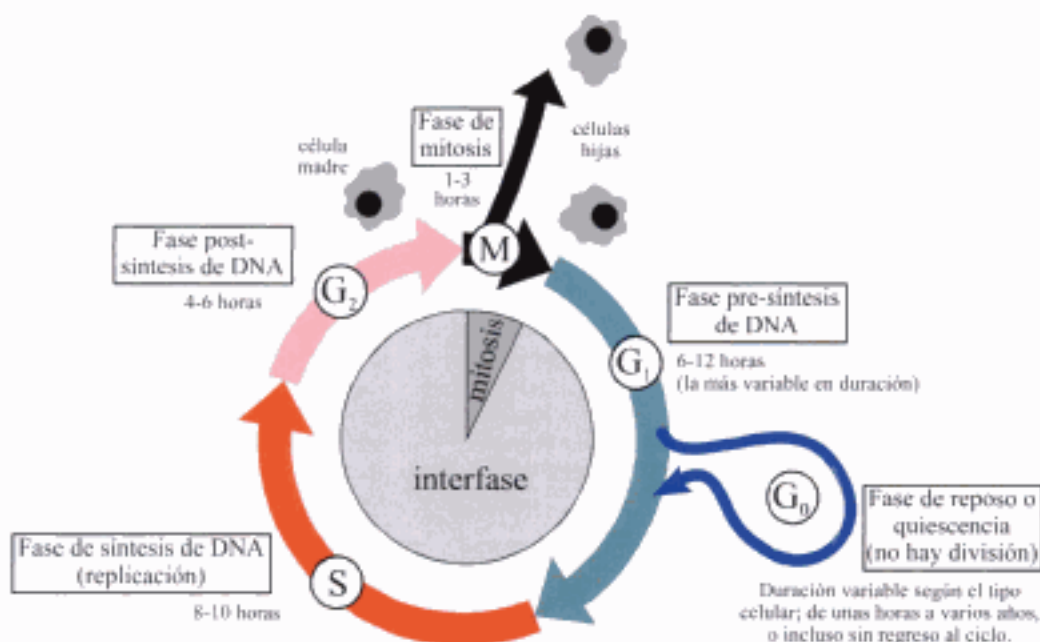
Mediante la división celular se generan tanto los gametos (división meiótica) como, a partir del cigoto, todas las células necesarias para que un organismo crezca y se desarrolle durante los periodos embrionario, fetal y de crecimiento hasta el estado adulto, y para la renovación continua de las células que mueren durante dichos periodos (división mitótica).

La división por mitosis tiene lugar normalmente durante toda la existencia del organismo, desde la división inicial del cigoto hasta la muerte del individuo. Se calcula que durante la vida humana se producen unas 10^{17} divisiones. La transición entre dos divisiones sucesivas constituye, junto a la propia división, el **ciclo de división celular** o simplemente **ciclo celular**. Su duración varía ampliamente de acuerdo con el tipo celular, así como en función de determinados estímulos. Algunas células evolucionan por una sucesión continua de ciclos de división (por ejemplo, células epiteliales del intestino y precursores hematopoyéticos en la médula ósea), otras poseen ciclos más espaciados (una vez al mes en células hepáticas, una vez cada 3 años en células endoteliales de vasos sanguíneos, etc.) y, como caso extremo, algunas células detienen su progresión en el ciclo, reanudándolo tras un periodo prolongado (quiescencia) o incluso no dividiéndose más (neuronas).

Algunas células se especializan en la producción de gametos o células sexuales, que permitirán la formación de un nuevo individuo. Tales células **germinales** (a diferencia del resto, células **somáticas**) se dividen mediante un mecanismo diferente, denominado **meiosis**, que también va precedido de un ciclo celular similar al de las células somáticas.

Una visión global del ciclo celular permite observar cómo la célula transcurre de manera ordenada y sin discontinuidad por dos grandes etapas bien diferenciadas: la **división celular** propiamente dicha (por mitosis para células somáticas y por meiosis para células germinales) y la **interfase** (común a ambas). Durante esta última, la célula se prepara para la siguiente división duplicando su material genético (DNA) y todo su contenido (proteínas, RNA, orgánulos, membranas), de modo que se duplica su tamaño antes de dividirse en dos células hijas. Clásicamente, el

ciclo se plantea en torno al proceso de replicación, o duplicación del DNA nuclear, con lo que se definen dentro de la interfase tres etapas: la fase S (durante la cual tiene lugar la replicación), la fase G_1 (previa) y la fase G_2 (posterior). Mientras que algunas células entran en división tras completar cada interfase, otras pasan a un estadio modificado de G_1 , llamado fase G_0 o quiescente, en el que permanecen días o incluso años hasta emprender un nuevo ciclo, o del que nunca salen (células que no se dividen, como es el caso de las neuronas).



Siguiendo el planteamiento general de este texto, interesa relacionar los aspectos genéticos clásicos del ciclo con aquellos otros de índole más bioquímica y molecular. Para ello, se abordarán en forma detallada la interfase y los dos tipos de división celular, con un múltiple objetivo:

- Poner de manifiesto la enorme variación estructural que supone, durante la interfase, la condensación de la molécula de DNA desde cromatina a cromosoma y su descondensación en sentido inverso.
- Resaltar el proceso de la replicación, que también tiene lugar durante la interfase, pero en un momento diferente del anterior. Su carácter molecular y enzimático se estudiará con detalle en el tema 12.
- Establecer las analogías y diferencias entre la división por mitosis de las células somáticas y la división por meiosis de las células germinales primarias para dar lugar a células germinales maduras o gametos.

9.2 LA CROMATINA EN LA INTERFASE

La **interfase**, o lapso que transcurre entre dos divisiones celulares, es normalmente el período más prolongado del ciclo celular, con sus tres fases G_1 , S y G_2 . Esta etapa termina cuando una célula somática entra en división por mitosis (para formar células somáticas hijas), y cuando una célula germinal primaria se divide por meiosis (para formar las células germinales maduras). Se va a estudiar cómo durante la interfase varía el grado de condensación del material genético y también cómo varía el contenido de DNA (2C en G_1 , 4C en G_2) sin modificarse el número de cromosomas (la célula se mantiene diploide, $2n$).

9.2.1 Variaciones en condensación y contenido en DNA de la cromatina

Durante la mayor parte de la vida celular (bien interfase o bien quiescencia, según los casos), la molécula de DNA no está presente en la forma extendida de doble hélice ni tampoco en la forma compacta de cromosomas, sino en el estado parcialmente condensado conocido como **cromatina** (págs. 85 y 87).

Al progresar en el ciclo celular, la cromatina se descondensa gradualmente durante la fase G_1 hasta adoptar una conformación **local** totalmente extendida (correspondiente a la doble hélice), necesaria para la separación de las dos hebras durante la replicación (fase S). Para ello, la eucromatina ha de pasar previamente por los estados de fibra de 30 y 10 nm (pág. 101).

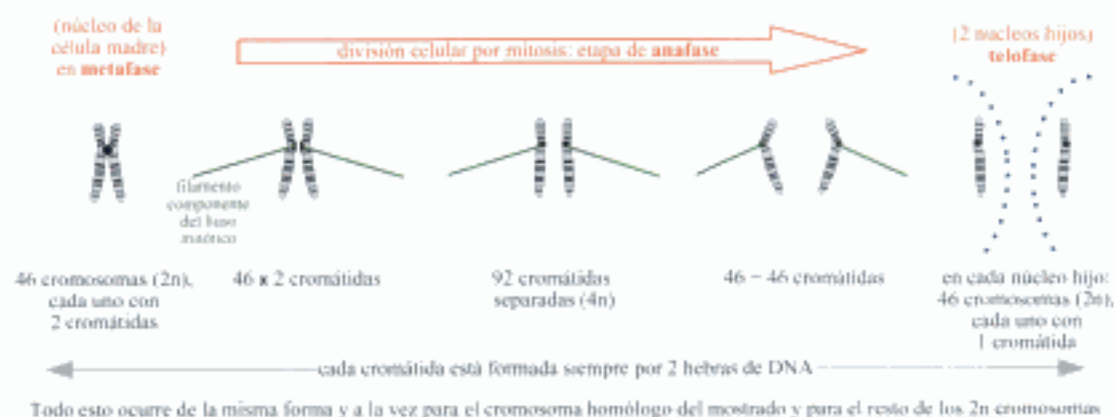
Durante la **fase G_1** , las células, tanto somáticas como germinales, son diploides ($2n$) y su contenido en DNA es 2C. En la **fase S o de síntesis** se produce la replicación del DNA de los $2n$ cromosomas individuales. Cada hebra de este DNA sirve de molde para la síntesis de otra nueva, que permanece asociada por apareamiento de bases. Las dos moléculas de DNA resultantes permanecen unidas por el **centrómero**, dando así lugar a cromosomas con 4 hebras de DNA; cada doble hebra constituye una **cromátida**. De este modo, el número de cromosomas permanece constante (diploide, $2n$), pero se duplica el contenido de DNA, desde 2C a 4C. Una vez finalizada la replicación, la célula entra en

la **fase G₂** donde se inicia la condensación gradual de la cromatina (pág. 100, en secuencia inversa a la descondensación ocurrida durante la fase G₁, pág. 101), que se completa en las primeras etapas de la mitosis dando lugar a cromosomas visibles al microscopio, con su aspecto típico de 2 cromátidas y 4 brazos. La célula sigue siendo 2n, 4C.

Aunque se puede considerar que cada molécula de DNA es un cromosoma en cualquier momento del ciclo, el término **cromosoma** corresponde estrictamente a esta forma condensada al máximo. A veces se usan los términos **cromosoma metafásico** y **cromosoma interfásico** para las formas condensada y descondensada (cromatina), respectivamente.

9.3 DIVISIÓN DE CÉLULAS SOMÁTICAS POR MITOSIS

La mitosis, o **fase M**, es el periodo más corto del ciclo celular, durante el que tiene lugar la división de la célula, que ya ha duplicado su material genético y citoplasmático, en dos células hijas idénticas. Clásicamente se considera dividida en cinco etapas: (1) **Profase**: la cromatina se condensa hasta formar los cromosomas, dejan de observarse los nucleolos y aparece el huso mitótico. (2) **Prometáfase**: comienza a romperse la membrana nuclear, se ven por primera vez las 2 cromátidas de cada cromosoma y los 2n cromosomas migran hacia el plano ecuatorial de la célula. (3) **Metafase**: los 2n cromosomas, totalmente condensados, aparecen alineados en el plano ecuatorial, independientes unos de otros. (4) **Anafase**: los cromosomas se dividen por sus centrómeros, separándose en dos cromátidas, que son desplazadas hacia polos opuestos de la célula. Cada cromátida es ya un cromosoma hijo independiente. (5) **Telofase**: comienzan a descondensarse los cromosomas y se forman membranas nucleares para delimitar dos núcleos. Paralelamente, la membrana celular separa el citoplasma en dos partes, cada una con un núcleo y la mitad de los orgánulos, formando las dos células hijas, en el proceso llamado **citocinesis** (o citoquinesis).



Como consecuencia de la mitosis, cada célula hija recibe una cromátida de cada cromosoma de la célula madre, es decir, la mitad de DNA, pero constituyendo un conjunto cromosómico completo (23 pares de cromosomas) con toda la información genética de la célula madre. Por ello, se dice que **la mitosis es un proceso de división no reductora** (pág. 104). La única diferencia es que la célula justo antes de dividirse tiene 2n cromosomas de 4 hebras y 2 cromátidas cada uno (4C) y tras la división cada célula hija tiene 2n cromosomas de 2 hebras y una cromátida cada uno (2C) (págs. 100,101).

9.4 DIVISIÓN DE CÉLULAS GERMINALES PRIMARIAS POR MEIOSIS

La división meiótica forma parte del proceso general de formación de **gametos** o células sexuales haploides, a partir de células germinales primordiales (que son diploides, al igual que las células somáticas), en el proceso conocido como **gametogénesis**.

9.4.1 Gametogénesis

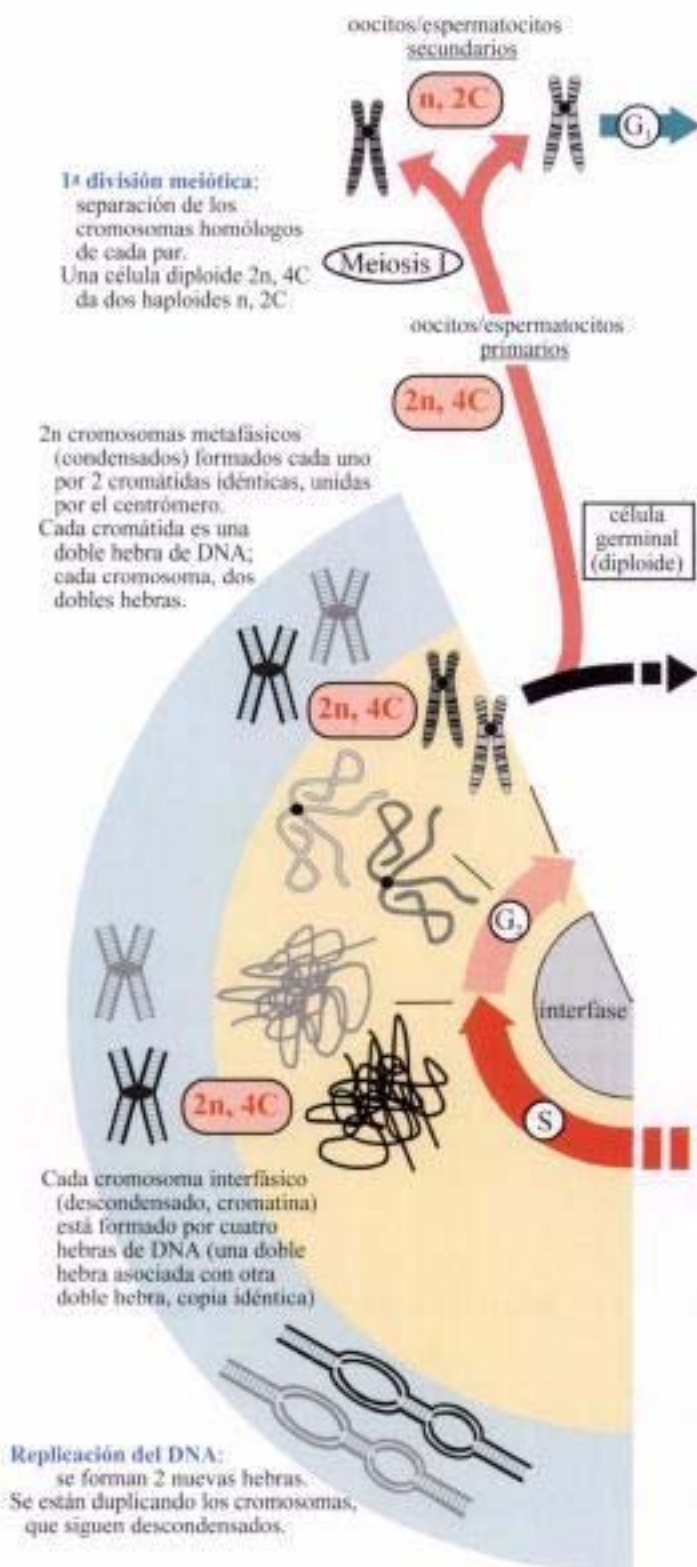
La ovogénesis y la espermatogénesis se inician a partir de **células germinales primordiales**, células especializadas del óvulo o testículo, mediante una serie de divisiones mitóticas que, con el consiguiente proceso de crecimiento y diferenciación celular, producen sucesivas generaciones de células llamadas **oogonios** y **espermatogonios** (también ogonia y espermatogonia, págs. 102,103). A continuación éstas, también por mitosis, dan lugar a **ocitos** y **espermatoцитos primarios**. Todas estas células son diploides. Finalmente, a partir de cada ocito o espermatoцитo primario y mediante dos divisiones sucesivas diferentes (meiosis I y meiosis II) se produce la reducción del estado diploide al haploide. Por ello, el conjunto de esas dos divisiones recibe el nombre de **meiosis** o **división celular reductora**.

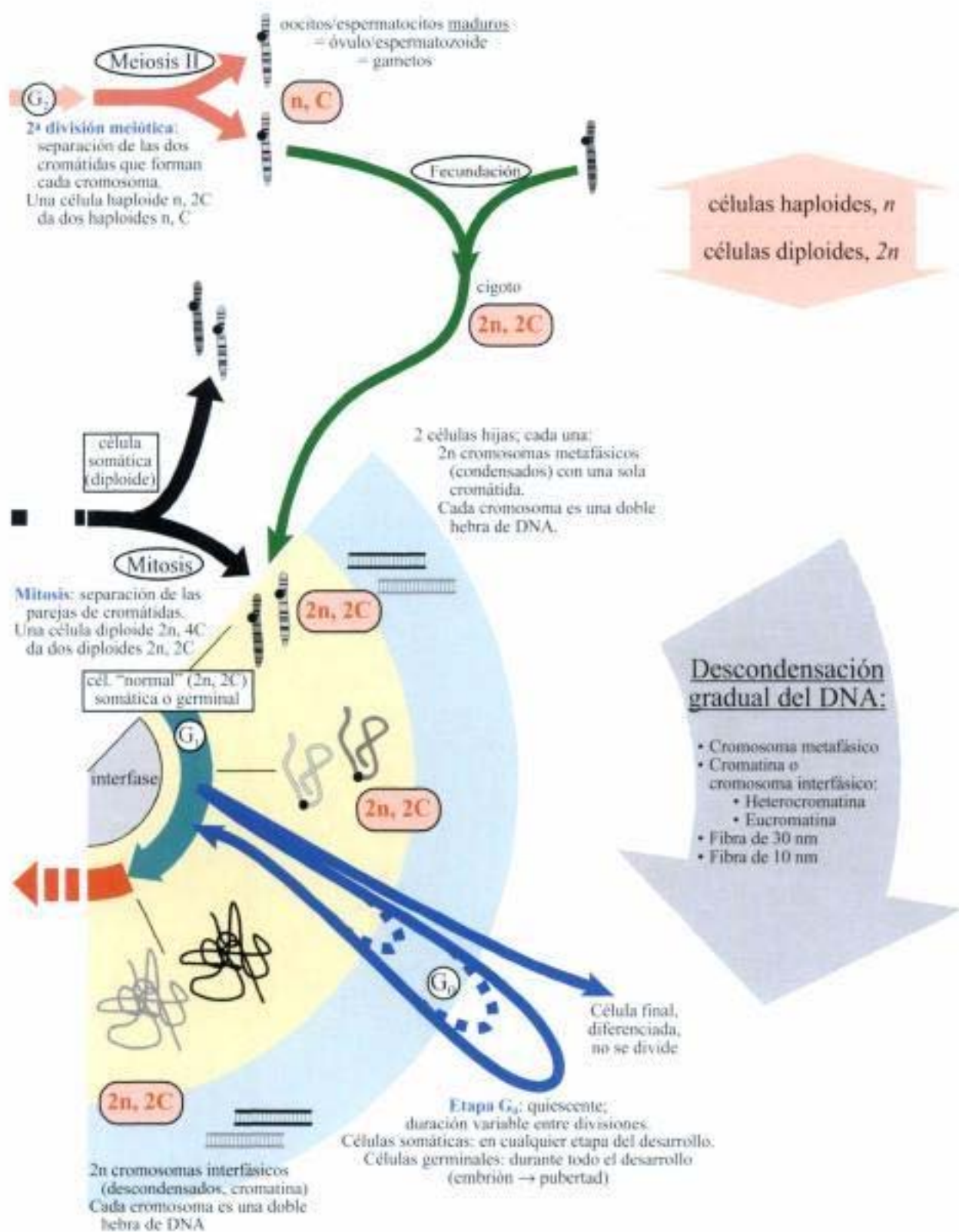
células haploides, n

células diploides, $2n$

Condensación gradual del DNA:

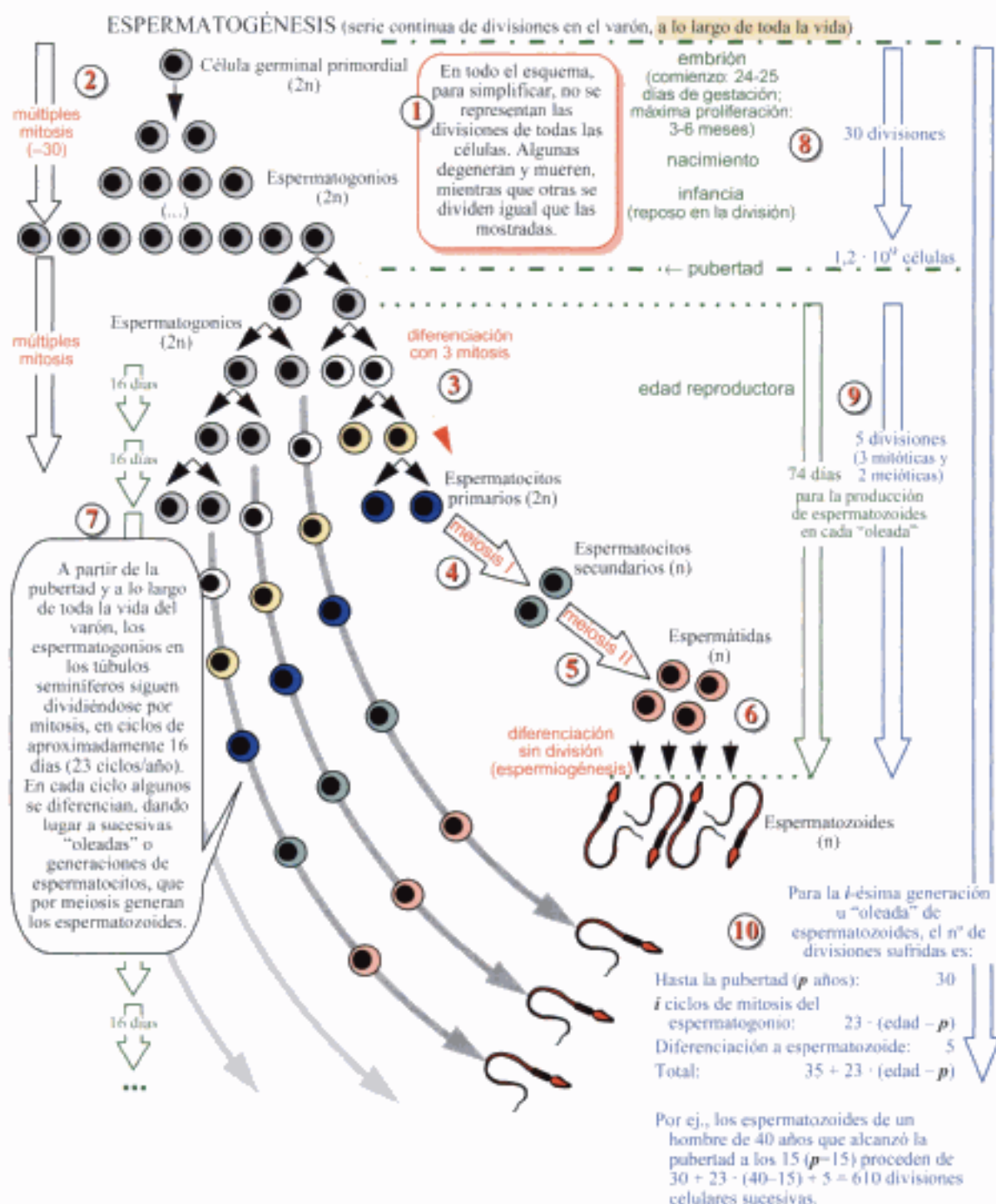
- Fibra de 10 nm
- Fibra de 30 nm
- Cromatina o cromosoma interfásico:
 - Eucromatina
 - Heterocromatina
- Cromosoma metafásico





9.4.1.1 Espermatogénesis

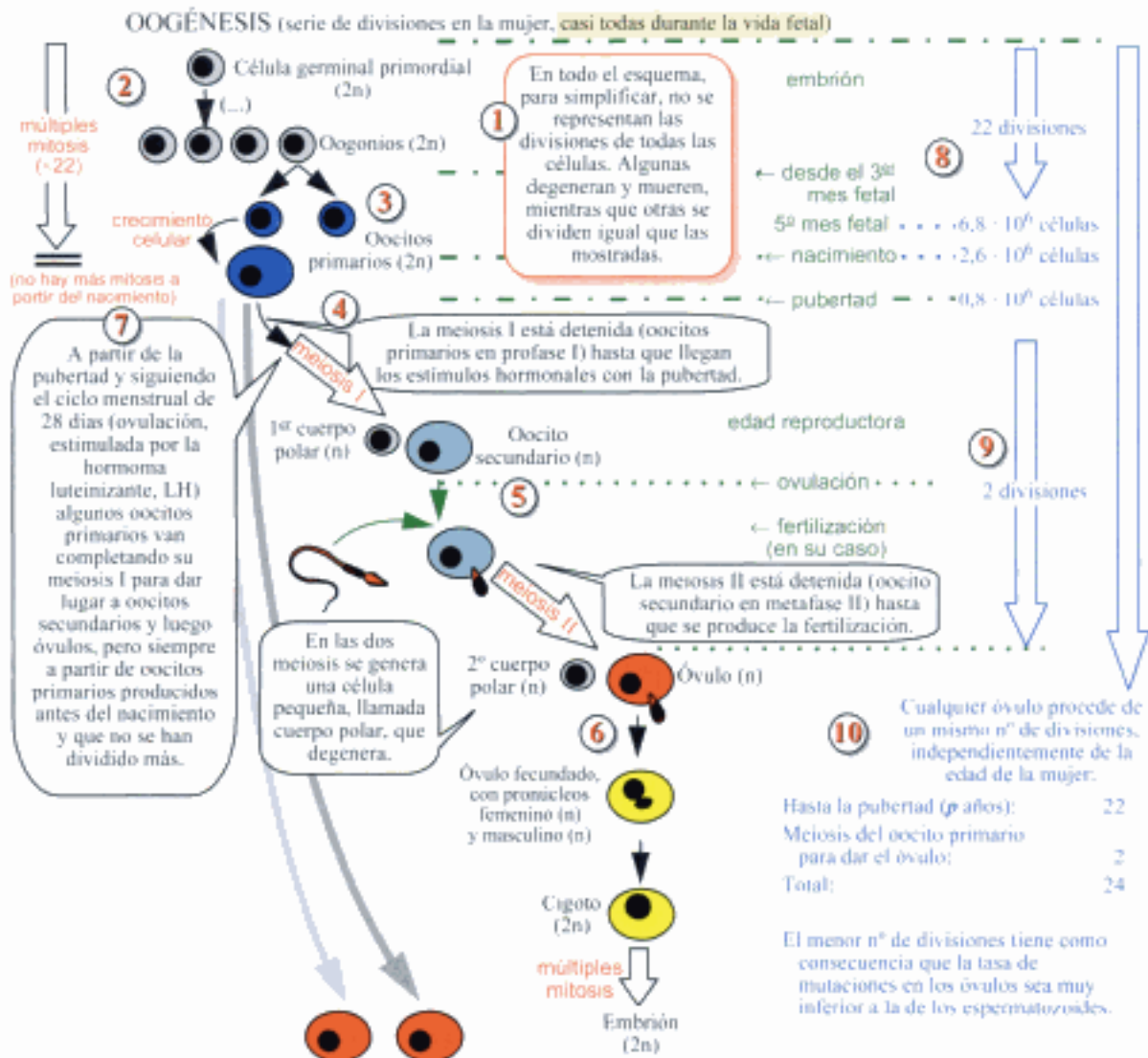
En el caso del varón, el proceso, que tiene lugar repetidamente a lo largo de toda su vida, culmina con la producción de espermatozoides o gametos masculinos.



Cada uno de los espermatocitos primarios ha dado lugar a un reparto simétrico del citoplasma en cuatro espermatozoides iguales. El número total de divisiones alcanza valores sorprendentes, variables según la edad, ya que dependen de los años transcurridos desde la pubertad. El número total de espermatozoides realmente producidos es obviamente variable; baste indicar que un varón sano puede formar unos 200 millones diarios durante toda su vida fértil y que un eyaculado, con unos 2 ml, contiene unos 60 millones de espermatozoides.

9.4.1.2 Oogénesis

Por el contrario, en la mujer la **oogénesis** ocurre en su mayor parte durante la vida fetal, a partir de la célula germinal primordial originada en el endodermo del saco vitelino del embrión. El mayor número de oocitos primarios se alcanza hacia el quinto mes de vida fetal (compárese con la pubertad para el varón) y es muy inferior al de espermatoцитos primarios. A partir de ese momento, el número de oocitos se reduce continuamente hasta la pubertad, se mantiene durante la vida fértil de la mujer y desaparece casi totalmente al llegar la menopausia.

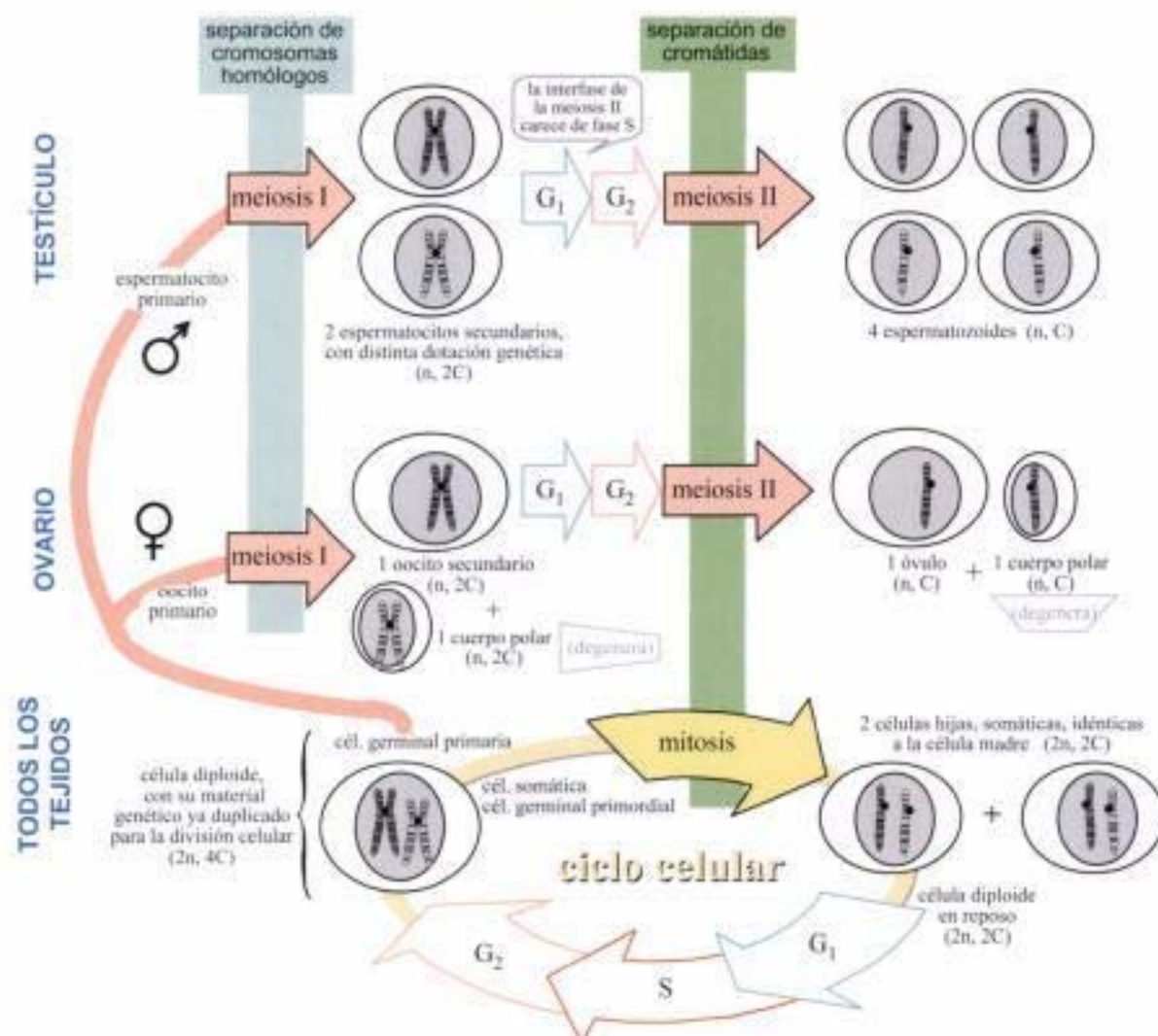


Otra peculiaridad que diferencia espermatogénesis y oogénesis es que en las dos divisiones meióticas de esta última el reparto del citoplasma es asimétrico: en la meiosis I se forma una célula pequeña llamada *cuerpo polar* y un oocito secundario grande. La división de éste en la meiosis II origina un *segundo cuerpo polar* y un óvulo maduro. Ambos cuerpos polares degeneran posteriormente. Por tanto, un oocito primario sólo genera un óvulo, mientras que un espermatoцитo primario produce cuatro espermatozoides.

9.4.2 Reparto del material genético en la meiosis

9.4.2.1 Carácter reductor de la meiosis

En la meiosis I, el gametocito (oocito o espermatocito) primario separa sus 23 pares de cromosomas homólogos ($2n$, $4C$, con dos cromátidas cada uno), de forma que pasan 23 cromosomas diferentes (uno de cada par; n , $2C$, con 2 cromátidas) a cada gametocito secundario (1ª Ley de Mendel). Se dice que se produce la *segregación* de los cromosomas, o que cada par ha sufrido *disyunción*; se reduce así a la mitad el número de cromosomas. Por ello, se dice que la meiosis es una división *reductora*, debido a la meiosis I. La meiosis II es *no reductora* y en cuanto al reparto de material genético se asemeja a la mitosis, como se discute a continuación.



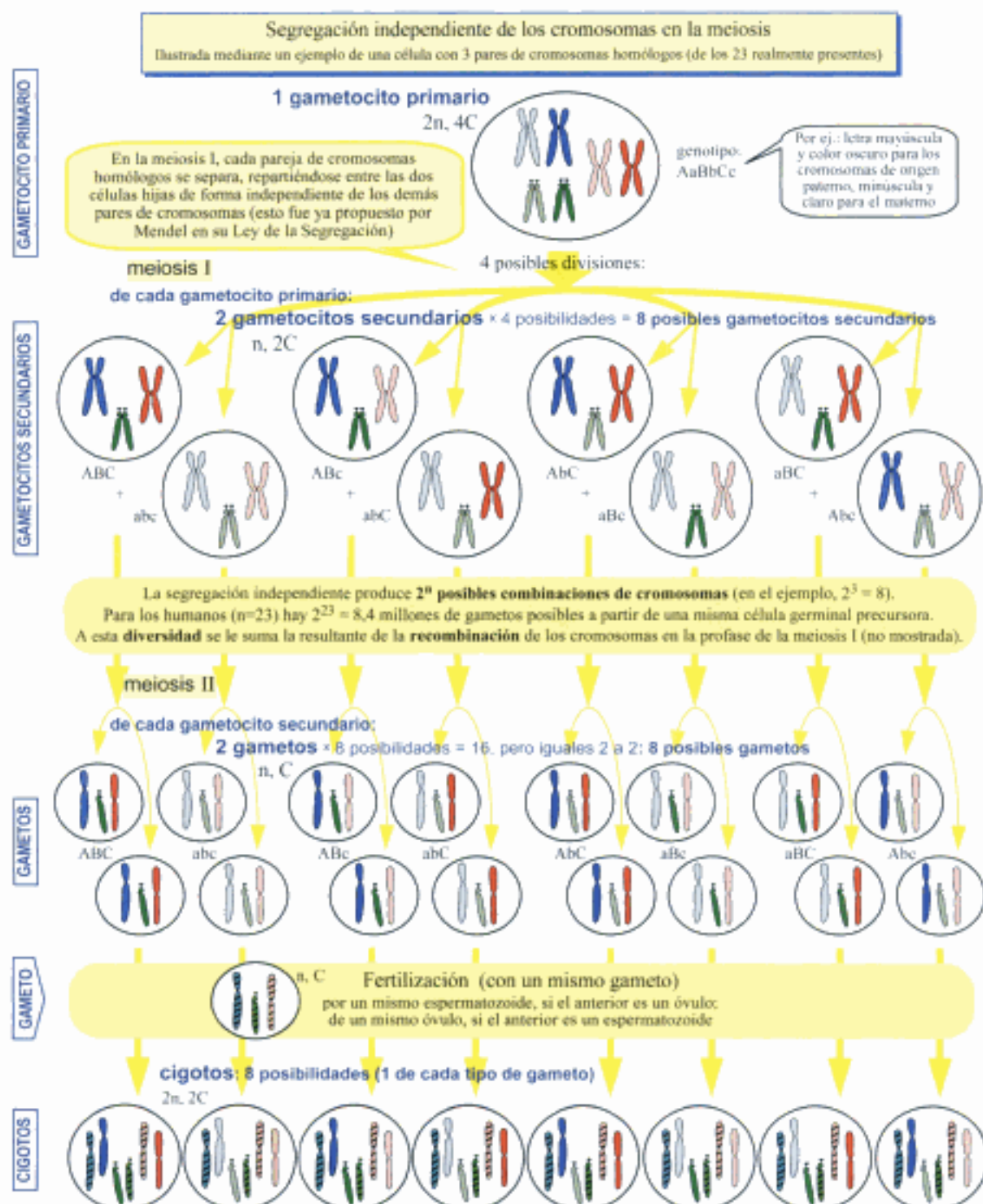
las 3 etapas de la interfase son previas tanto a mitosis como a meiosis I

9.4.2.2 Comparación entre meiosis y mitosis

La diferencia esencial entre la mitosis y las dos etapas de la meiosis radica en la distribución del material genético entre las células hijas (somáticas diploides en mitosis; gametos haploides en meiosis). El mecanismo de la meiosis II es equivalente al de la mitosis, puesto que **se separan las dos cromátidas** y se reduce la cantidad de DNA ($2C \rightarrow C$) sin cambiar el carácter haploide (n). De este modo, cada célula hija recibe **copias idénticas** de DNA (dobles hebras duplicadas en la replicación). En contraste, en la meiosis I **se separan los miembros de cada pareja de cromosomas**, con lo que se reducen a la vez la cantidad de DNA ($4C \rightarrow 2C$) y el número de cromosomas ($2n \rightarrow n$), y las células hijas reciben **diferente material genético** (pues cada cromosoma de un par puede contener alelos diferentes en cada locus, pág. 94).

9.4.2.3 Segregación de cromosomas y diversidad genética

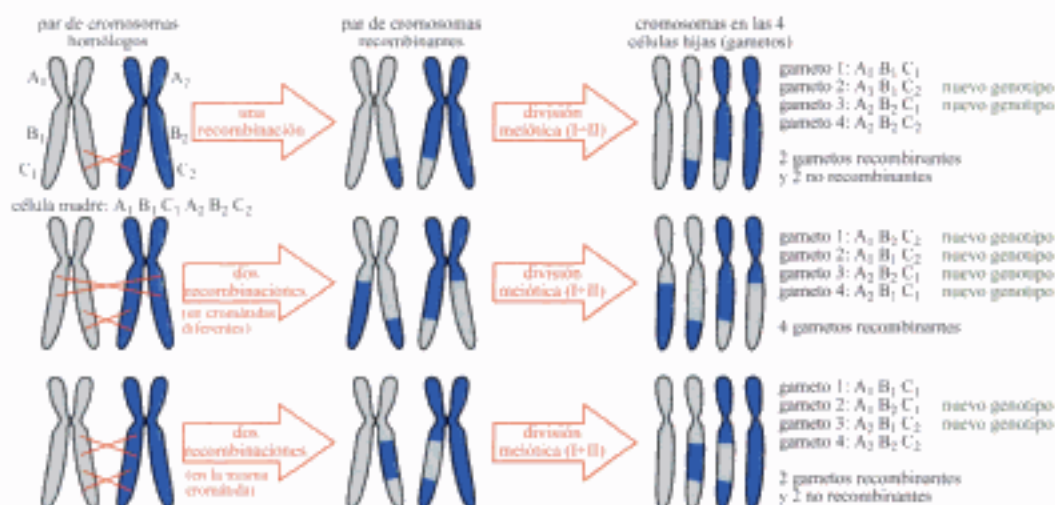
En **ambas meiosis**, el reparto de los cromosomas homólogos de cada par (meiosis I) y de las cromátidas de cada cromosoma (meiosis II) se hace al azar, de forma totalmente independiente para cada cromosoma (2ª Ley de Mendel). En definitiva, se asegura con creces que cada gameto recibe diferente material genético; esta variabilidad es el fundamento del polimorfismo genético, cuyos detalles y consecuencias se estudian más adelante (tema 26).



Estos 8 cigotos proceden de la fertilización por un mismo gameto. Si se considera el total de posibilidades, de unos mismos padre y madre hay 8 posibles gametos masculinos y 8 posibles gametos femeninos, lo que da un total de 64 combinaciones en el cigoto. Por tanto, en humanos, 2^{23} posibles gametos masculinos y 2^{23} posibles gametos femeninos dan 70.4 billones de posibles cigotos.

9.4.2.4 Recombinación meiótica

Durante la profase de la meiosis I (profase I) tiene lugar un proceso de gran trascendencia en cuanto a la herencia y a la generación de la diversidad genética de cada individuo: se trata de la **recombinación meiótica**, o simplemente **recombinación**. Los detalles de este proceso no se tratan aquí, pero básicamente consiste en el **sobrecruzamiento**, o intercambio de regiones de DNA entre cromátidas de los dos cromosomas homólogos (es decir, nunca hay recombinación entre cromátidas de distinto par). De ahí que se llame también **recombinación homóloga**.

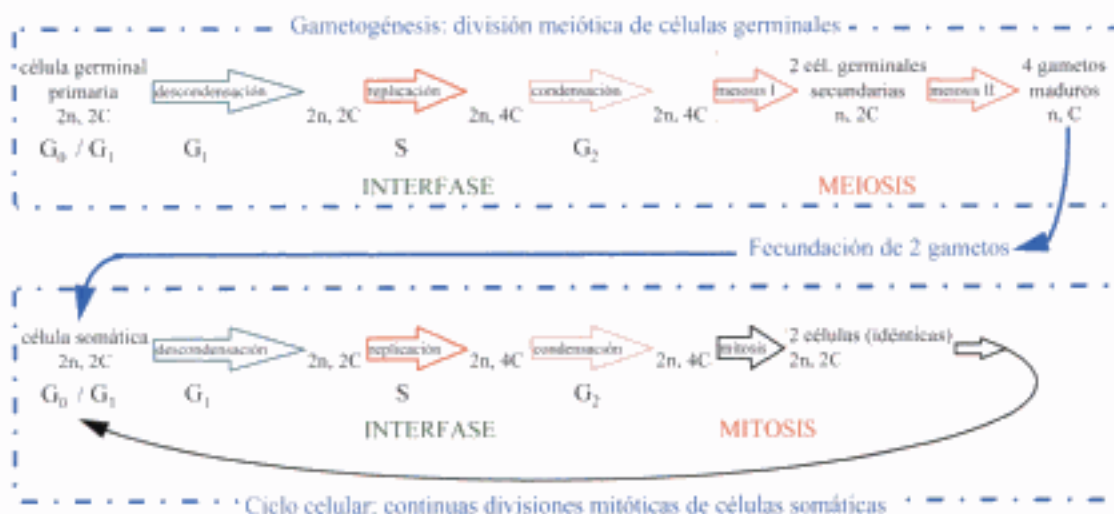


9.5 UNIÓN DE LOS GAMETOS HAPLOIDES DURANTE LA FECUNDACIÓN

Cuando los dos gametos (n , C) se unen en la **fecundación**, el pronúcleo macho del espermatozoide se funde con el pronúcleo hembra del óvulo, combinando sus dos genomas haploides para dar un genoma $2n$, $2C$ en la célula huevo o cigoto. Ésta es así la primera célula somática propiamente dicha. Como tal, entra en la fase G_1 y posteriores de la interfase. A partir del cigoto se forman en definitiva todas las células somáticas. El mecanismo que rige la transformación de una simple célula, el cigoto, en un individuo constituye un tema de gran interés biológico, casi totalmente desconocido desde el punto de vista molecular. Tras múltiples divisiones celulares, se origina un amasijo globular de células embrionarias, las cuales se autoorganizan y especializan hasta formar un fantástico conjunto de células (220.000 millones), subdivididas en casi 250 familias diferentes, cada una con una función y disposición concretas, en las distintas estructuras (cerebro, ojos, nariz, corazón, etc.) del organismo completo. La Biología del Desarrollo es un campo de activa investigación y en los últimos años se han aislado genes que controlan, por ejemplo, la disposición dorsal/ventral o derecha/izquierda de algunos órganos; sin embargo, aún estamos lejos de conocer los mecanismos moleculares que gobiernan la organogénesis.

9.6 RESUMEN INTEGRADO

Para terminar, el siguiente esquema resume los distintos procesos estudiados, recogiendo de una forma unificada los cambios en la cantidad de material genético que sufre la célula durante los procesos de fecundación, división meiótica y mitótica y ciclo celular.



Tema 10

Organización del genoma eucariótico

10.1	Complejidad del genoma eucariótico	107
10.2	DNA de copia única, simple o no repetitivo	108
10.3	DNA repetitivo	109
10.3.1	DNA repetitivo codificante	110
10.3.1.1	DNA repetitivo codificante agrupado	110
a)	Familias génicas clásicas	110
b)	Familias multigénicas de genes agrupados	111
10.3.1.2	DNA repetitivo codificante disperso: familias multigénicas de genes dispersos	112
10.3.2	DNA repetitivo no codificante	113
10.3.2.1	DNA altamente repetitivo y agrupado: DNA satélite	113
10.3.2.2	DNA moderadamente repetitivo y disperso	114
a)	Bloques dispersos de repeticiones en tándem: DNA minisatélite y DNA microsatélite	114
b)	Repeticiones dispersas: SINEs y LINEs	114
10.4	Estudio experimental de la complejidad en el genoma de eucariotas	115

En **procariotas**, prácticamente todo el DNA existe en forma de copia única y codifica productos génicos (proteínas y RNAs). Sin embargo, esto no se cumple en el genoma nuclear de **eucariotas**; por ejemplo, se calcula que sólo 100 Mb del total de 3.300 Mb del genoma humano haploide son codificantes. Además, el genoma eucariótico se caracteriza por la repetición de secuencias. El objetivo de este tema es, por tanto, plantear esta distinción entre *DNA de copia única* y *DNA repetitivo* y estudiar con detalle este último en el genoma humano. El carácter *codificante*, aunque se incluye también aquí, se comprenderá mejor en temas posteriores, al estudiar la expresión génica (transcripción, maduración y traducción). Los aspectos experimentales que han permitido la demostración de la presencia de DNA repetitivo en el genoma eucariótico se describen al final del capítulo. Basándose en toda esta información, se podrá abordar más adelante su aplicación a problemas biotecnológicos en general y clínicos en particular.

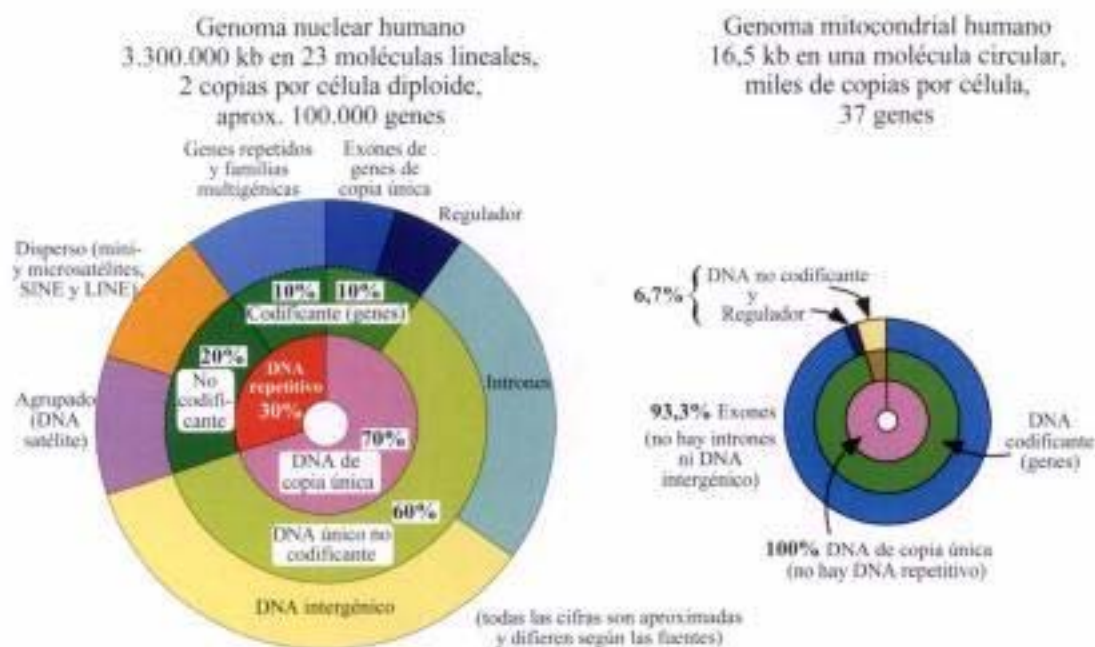
10.1 COMPLEJIDAD DEL GENOMA EUCARIÓTICO

La presencia de DNA repetitivo da lugar al concepto de "**complejidad**" del genoma. Este término se define como la suma de tamaños (pb) de todas sus secuencias diferentes (es decir, las secuencias repetidas se contabilizan sólo una vez). En consecuencia, para un tamaño de genoma dado, a mayor grado de repetición, menor "complejidad".

Nota: La interpretación de este término se puede prestar a confusión, pues frecuentemente se usa con la acepción común de la palabra complejidad, que puede parecer contraria a la definición anterior. Así, un genoma con gran número de secuencias repetidas es aparentemente más complejo, pero por definición su "complejidad" es menor. Para evitar confusiones, se usará entrecorillado cuando se refiera a la definición específica dada.

Para el estudio de la "complejidad" del genoma humano, éste puede dividirse en distintas categorías de DNA en función de su repetitividad y su carácter codificante:

DNA de copia única	Codificante: genes	Estructural: exones (codifica proteínas o RNAs)
	No codificante (función desconocida)	Regulador (controla la expresión del gen)
DNA repetitivo	Codificante	Intrones o DNA intragénico (intercalado dentro de los genes)
		DNA intergénico (separa los genes)
	No codificante	Agrupado: Genes repetidos en tándem Familias multigénicas agrupadas
		Disperso: Familias multigénicas dispersas
	Agrupado (en regiones heterocromáticas) y Altamente repetitivo	Repeticiones en tándem (DNA satélite)
		Bloques dispersos de repeticiones en tándem: Minisatélites, Microsatélites
	Disperso (por todo el genoma) y Moderadamente repetitivo	Repeticiones dispersas: SINE: secuencias <i>Alu</i> y otras, LINE: secuencias <i>Kpn</i> y otras



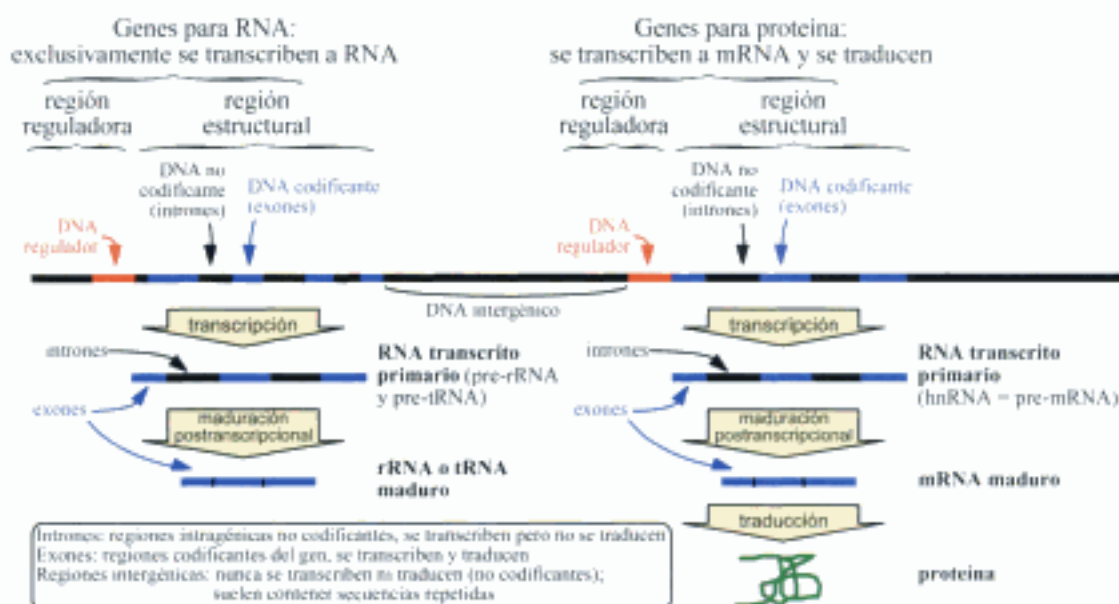
A diferencia del nuclear, el genoma de la mitocondria, de tamaño muy pequeño (16.569 pb en humanos, 200.000 veces inferior al genoma nuclear haploide), está formado por DNA codificante casi en su totalidad.

10.2 DNA DE COPIA ÚNICA, SIMPLE O NO REPETITIVO

Constituye la mayor parte del genoma, aunque en una proporción variada dependiendo del tipo de organismo (100% en procariotas, 80% en eucariotas inferiores, 50-70% en animales superiores). Parte de este DNA (~5%) constituye las secuencias de genes, que codifican los RNAs y proteínas celulares; otra parte (~5%) es responsable del control de la expresión de esas secuencias, mientras que el resto, mayoritario (~60%), es DNA no codificante, cuya función, si existe, apenas se conoce.

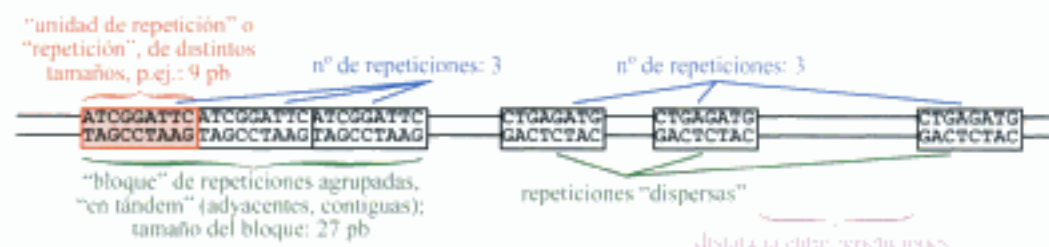
El DNA de copia sencilla fue considerado durante mucho tiempo como responsable único del concepto clásico de gen; hoy se sabe que también una pequeña parte del DNA repetitivo forma algunos genes eucarióticos, como se estudia

a continuación. El concepto de gen, aunque ha sufrido una gran evolución en eucariotas, para hacerse más complicado y con múltiples componentes, se sigue definiendo bajo un criterio estrictamente funcional, siempre en relación con la transmisión de la información desde DNA a RNA y desde mRNA a proteínas. Concretamente, un gen es aquella región de DNA que comprende, por un lado, las **secuencias estructurales**, que se transcriben exclusivamente en RNA (región estructural de genes de tRNA y rRNA) o bien se transcriben en un mRNA que a su vez da lugar a una proteína, y por el otro las **secuencias reguladoras**, que actúan en ambos casos controlando la transcripción de las regiones estructurales. El estudio detallado de las características de los genes eucarióticos se deja para temas posteriores, cuando se analice su expresión.



10.3 DNA REPETITIVO

Las secuencias de DNA que aparecen en copias múltiples, o DNA repetitivo, constituyen entre un 20 y un 50% del total del genoma humano, según autores. Esta abundancia es la característica determinante de la enorme "complejidad" del DNA eucariótico. Las secuencias repetidas, llamadas **unidades de repetición** o simplemente **repeticiones**, tienen tamaños diversos y cada una se encuentra de forma idéntica o casi idéntica muchas veces en el genoma. Esta distribución puede ser en **forma dispersa** por todo el genoma, entremezcladas con las secuencias de copia única, o bien en **forma agrupada**, localizada en regiones concretas del cromosoma. En ambos casos, el número de copias de la unidad de repetición varía desde unos cientos o miles (**DNA moderadamente repetitivo**) hasta cientos de miles (**DNA altamente repetitivo**). Esto coincide en muchos casos con una distribución en forma dispersa y agrupada, respectivamente.



Una parte del DNA repetitivo tiene carácter **codificante**, es decir, contiene la información para expresar un producto funcional (RNA o proteína). Para el resto de DNA repetitivo no se conoce una función clara; alguno posiblemente contribuya a mantener la estructura de los cromosomas, mientras que se ha llegado a proponer que una parte sea **DNA chatarra** o **DNA basura**, un vestigio evolutivo sin función actual. Finalmente, debe destacarse que algunas de las secuencias repetidas no codificantes (concretamente, las denominadas **minisatélites** y **microsatélites**), aun con función desconocida, tienen gran relevancia aplicada en pruebas de identificación y estudios familiares, debido a su variabilidad entre individuos.

La repetitividad del DNA tiene probablemente un origen evolutivo. Por un lado, las repeticiones agrupadas pueden surgir por errores en la replicación o en la recombinación genética. Por otro, la separación entre repeticiones dispersas puede proceder de translocaciones o transposiciones cromosómicas.

10.3.1 DNA repetitivo codificante

Se le llama también **DNA repetitivo funcional**, pues forma genes que se expresan (aunque no todos, como se verá). Su tamaño se estima entre el 10 y el 15% del genoma. Aparece en forma de "familias de genes", cuyos miembros se caracterizan por su homología, al haberse originado mediante duplicaciones (pág. 367) y variaciones de un gen ancestral.

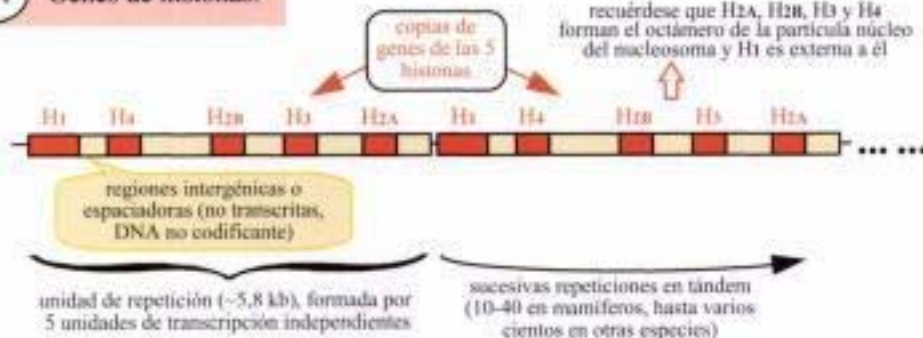
Tipos de DNA repetitivo codificante y ejemplos característicos:

1. Agrupado:		
a) Familias génicas clásicas o conservadas, con genes repetidos en tándem Muestran un alto grado de homología de secuencia, es decir, son copias prácticamente idénticas, lo que indica su relación evolutiva y funcional Se expresan todas las repeticiones	Histonas: Ej. 1 rRNAs: Ej. 2 tRNAs snRNAs	
b) Familias multigénicas con genes agrupados Situados en localizaciones específicas del genoma (pero sin relación con las regiones centroméricas y teloméricas) Sus repeticiones presentan menor homología: variantes, pseudogenes, genes truncados y fragmentos de genes Sólo se expresan algunas de las repeticiones	Globinas: Ej. 3 HLA-I (y otras muchas)	
2. Disperso: Familias multigénicas con genes dispersos		
Normalmente, cada familia está formada por un número pequeño de repeticiones repartidas por todo el genoma (incluso en cromosomas distintos), que casi siempre son todas funcionales		Aldolasa Actina Ferritina GA3PDH ...

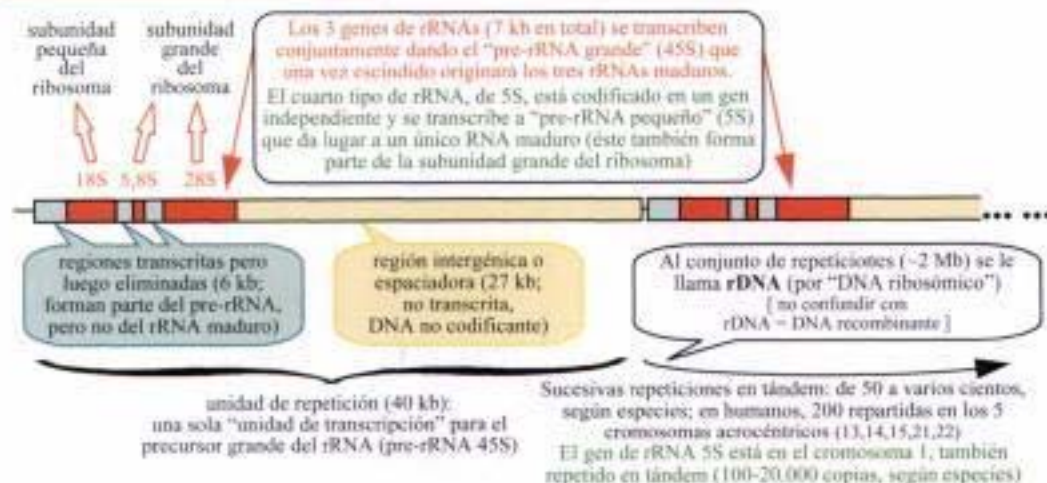
10.3.1.1 DNA repetitivo codificante agrupado

a) Familias génicas clásicas

Ej.1 Genes de histonas:



Ej.2 Genes de RNA ribosómico:



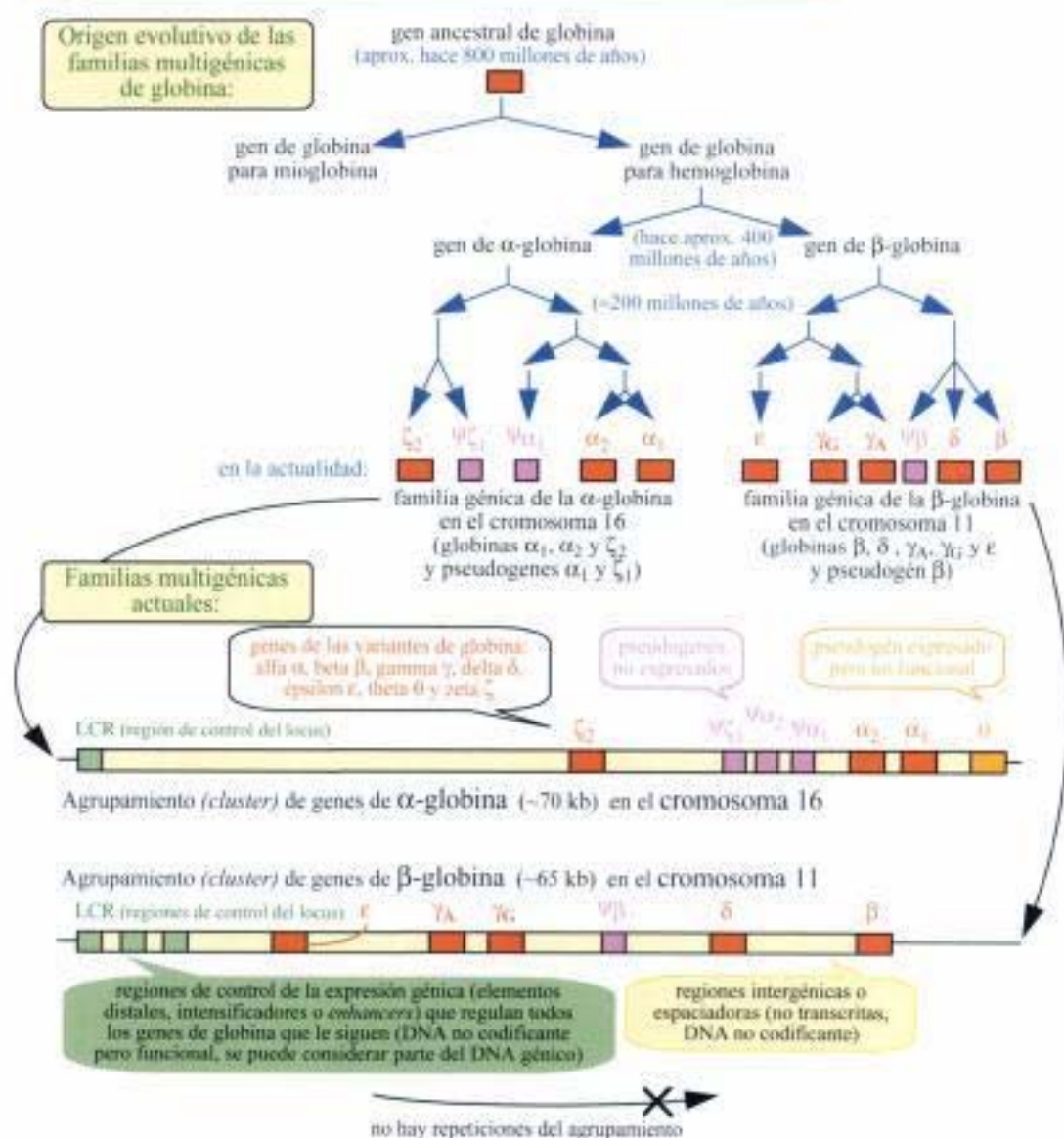
Contienen secuencias repetidas que codifican un mismo RNA o proteína, que puede así sintetizarse con rapidez y en gran cantidad. Son ejemplos característicos los genes de histonas y los de RNA ribosómico, productos ambos que la célula necesita en gran cantidad (recuérdese que las histonas deben sintetizarse rápidamente en la fase S del ciclo celular, para asociarse al nuevo DNA procedente de la replicación, y el rRNA supone el 75% del total de RNA celular).

b) Familias multigénicas de genes agrupados

A diferencia de las anteriores, los miembros de cada familia son genes con mayor diversidad de secuencias. Se pueden incluir en este grupo tres tipos de familias:

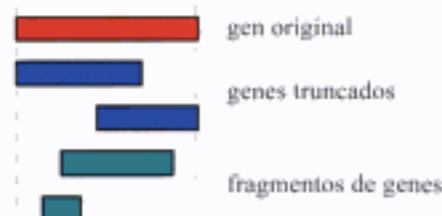
- Aquéllas con homología sólo en parte de la secuencia de DNA, que da lugar a proteínas con grandes regiones de secuencia y estructura comunes (dominios).
- Familias con homología escasa en las secuencias del DNA, lo que genera proteínas con pequeños motivos comunes, cada uno de pocos aminoácidos.
- Grandes "superfamilias", con muy débil homología en la secuencia de DNA y, por tanto, en la de las proteínas, a pesar de lo cual éstas mantienen una similitud estructural (en parte de su molécula, dominios comunes) y una relación funcional. Los genes que forman cada superfamilia parecen estar relacionados evolutivamente, pero más distantes entre sí que los miembros componentes de las familias anteriores.

Ej.3 Familias de genes humanos de las cadenas α y β de globina :



Como consecuencia de los procesos evolutivos que dan lugar a familias multigénicas, aparecen con frecuencia en ellas miembros con características diferenciales:

- **Copias de genes (variantes)** con pequeñas diferencias de secuencia, que codifican productos génicos funcionales pero de características ligeramente distintas (por ejemplo, diferencias en actividad enzimática, parámetros cinéticos, especificidad por sustrato, afinidad por su ligando, etc.). Es común que se active de forma alternativa la expresión de una u otra variante en función del estado de desarrollo, de diferenciación celular, tipo de tejido ...
- **Pseudogenes:** se aplica este término (abreviado como Ψ) a copias inactivas de un gen, es decir, que nunca se expresan, o que aun expresándose no son funcionales. Generalmente, la inactivación se ha producido por acumulación de mutaciones durante la evolución en lo que inicialmente era un duplicado del gen. Existen tipos diversos de pseudogenes, en función de si se expresan, si sufren procesamiento postranscripcional, así como de su origen evolutivo.
- **Genes truncados:** se trata de copias incompletas del gen, por pérdida de una región situada en uno de los extremos (5' o 3').
- **Fragmentos de genes:** en este caso se conserva sólo una porción del interior de la secuencia del gen (se han perdido ambos extremos).



Obviamente, sólo las copias idénticas (familias clásicas) y las variantes conducen a productos génicos funcionales, mientras que los pseudogenes, genes truncados y fragmentos de genes parecen ser vestigios evolutivos sin función actual.

Los genes de la globina constituyen uno de los ejemplos más representativos de familias multigénicas con genes agrupados, implicada en una función tan importante como es el transporte de oxígeno en la corriente sanguínea. Su estudio implica consideraciones evolutivas, de desarrollo y diferenciación, regulación de la expresión génica ...

En cada etapa del desarrollo humano es un órgano diferente el único que expresa los genes de globina. Además, en cada órgano se activan genes diferentes

Las "regiones de control del locus" (**LCR**) de ambas familias multigénicas **activan la expresión** de uno u otro de los genes dependiendo de la etapa de desarrollo y órgano

Cada etapa del desarrollo se caracteriza, por tanto, por la presencia de diferentes hemoglobinas.
 P.ej., HbF (fetal) = $\alpha_2\gamma_2$
 HbA (adulto) = $\alpha_2\beta_2$

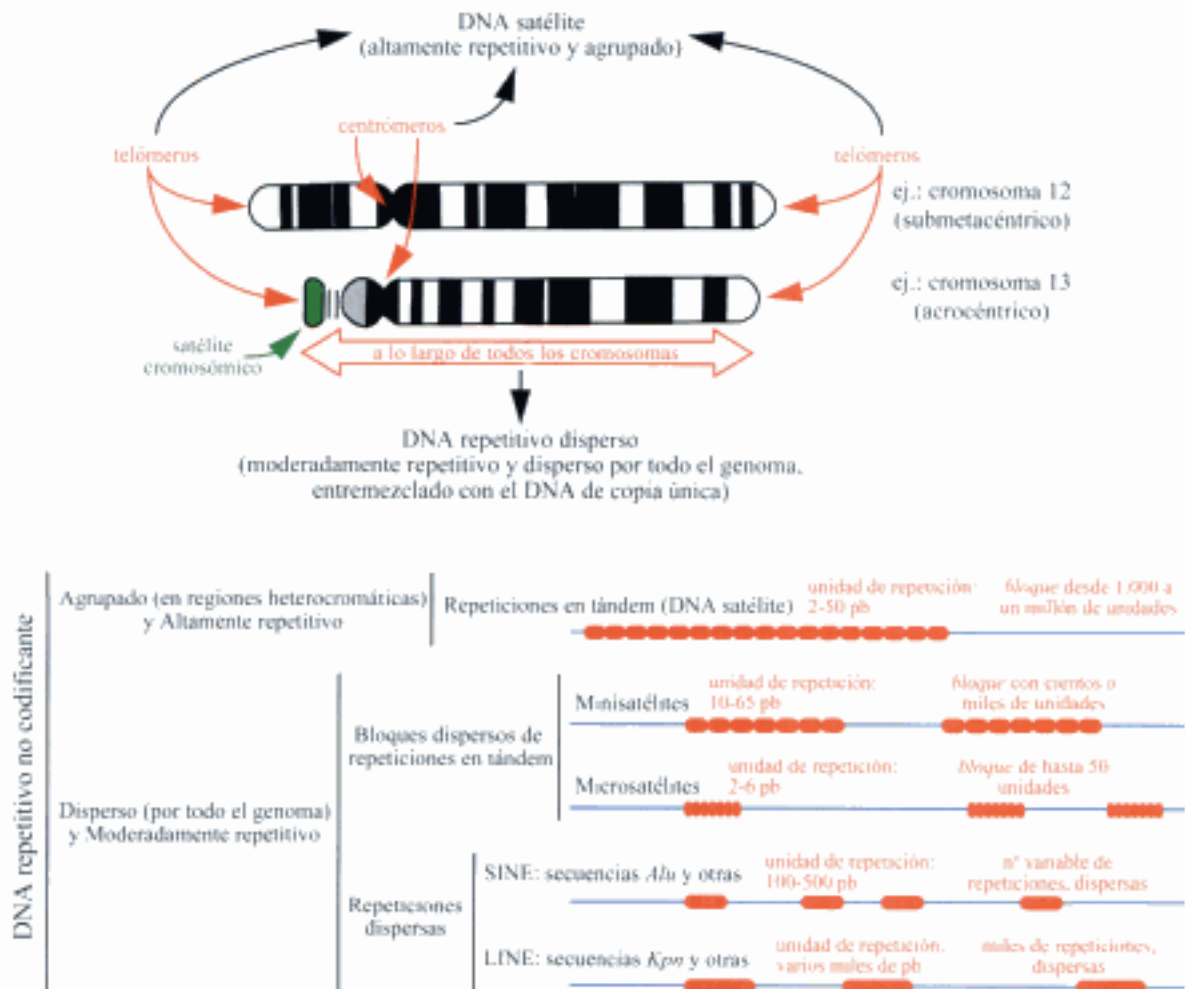
Etapa de desarrollo y órgano donde se sintetizan las globinas	Genes que se expresan en la familia génica de la α -globina (en el cromosoma 16)	Genes que se expresan en la familia génica de la β -globina (en el cromosoma 11)	Hemoglobina(s) sintetizada(s)
Embrión (saco vitelino)	 Mayoritaria: globina ζ (zeta)	 Mayoritaria: globina ϵ Minoritaria: globina γ	 Mayoritaria: $\zeta_2\epsilon_2$ Minoritaria: $\zeta_2\gamma_2$
Feto (hígado y bazo)	 Mayoritaria: globina α	 Mayoritaria: globina γ Minoritaria: globina β	 Mayoritaria: $\alpha_2\gamma_2$ Minoritaria: $\alpha_2\beta_2$
Adulto (médula ósea)	 Mayoritaria: globina α	 Mayoritaria: globina β (97%) Minoritaria: globina δ (2%)	 Mayoritaria: $\alpha_2\beta_2$ Minoritaria: $\alpha_2\delta_2$

10.3.1.2 DNA repetitivo codificante disperso: familias multigénicas de genes dispersos

Mientras que en las familias multigénicas anteriores las repeticiones se disponen **agrupadas**, próximas en la secuencia lineal del genoma, en otras los genes que forman la familia (copias idénticas, variantes, pseudogenes ...) se localizan de forma **dispersa** por el genoma, a menudo incluso en cromosomas distintos. Entre estas familias se encuentran genes que codifican proteínas de funciones diversas: enzimáticas, reguladoras, de almacenamiento, estructurales, etc. Por ejemplo, la familia de aldolasa comprende 3 genes funcionales y 2 pseudogenes, repartidos en 5 cromosomas; los genes de actina forman otra familia con 4 genes y al menos 16 pseudogenes.

10.3.2 DNA repetitivo no codificante

Esta es la fracción del genoma a la que el término *DNA repetitivo* se refiere de forma más habitual. Su tamaño se estima entre un 20 y un 40% del DNA genómico total. Como ya se ha indicado, se subdivide en dos categorías: el moderadamente repetitivo, que se ubica en forma dispersa, y el altamente repetitivo, que además es agrupado. Su distribución general en el genoma puede representarse gráficamente como se muestra en la figura, para dos cromosomas representativos.

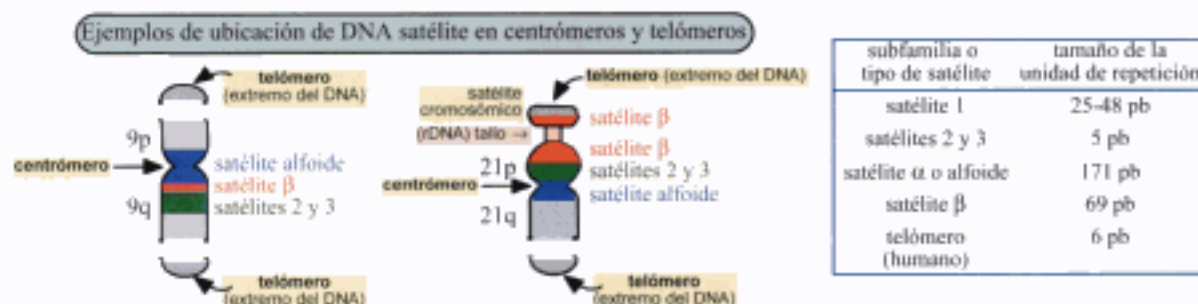


10.3.2.1 DNA altamente repetitivo y agrupado: DNA satélite

Supone entre un 10 y un 30% del genoma, y se le llama también **DNA repetitivo en tándem**. No debe confundirse con el DNA *minisatélite* y *microsatélite* (ambos corresponden al DNA moderadamente repetitivo y disperso, pág. 114) ni con los *satélites cromosómicos*, las pequeñas masas de cromatina que forman el extremo de los brazos cortos de un cromosoma acrocéntrico (pág. 88).

Este tipo de DNA está formado por unidades de pequeño tamaño (2-50 pb) repetidas en tándem entre miles y un millón de veces, con lo que el bloque de repetición abarca cientos de kb o incluso varias Mb. También es característica su **localización agrupada** en lugares específicos de varios cromosomas, correspondientes a las **regiones heterocromáticas**, principalmente en torno al centrómero y en los telómeros. Se conoce el papel de este DNA satélite en las regiones teloméricas (págs. 92 y 157), pero aún no su función exacta en el centrómero.

El nombre de *DNA satélite* tiene un origen metodológico: debido a su alta tasa de repetición, las regiones de este tipo de DNA poseen una composición de bases diferente a la del resto del genoma, que es más promediada. Cuando el DNA total se fragmenta (por ejemplo, por fuerzas de cizalla) y se analiza en gradientes de densidad de CsCl (pág. 137), los fragmentos procedentes de las regiones de DNA satélite, que poseen menor densidad debido a su menor contenido en G+C, se separan de la banda principal de DNA formando "bandas satélites". Cada una de ellas corresponde a una subfamilia de DNA satélite (tales como 1, 2, 3, α y β), con distinta secuencia de repetición y/o número de repeticiones.



10.3.2.2 DNA moderadamente repetitivo y disperso

Como su nombre indica, este tipo de DNA se distribuye a lo largo de todos los cromosomas, con un número de repeticiones no muy elevado (entre 10^2 y 10^4). Se subdivide en dos categorías, en función de la forma como se distribuyen las repeticiones: en primer lugar, algunas secuencias se repiten en tándem formando bloques, los cuales aparecen de forma dispersa por el genoma. El segundo tipo consiste en repeticiones igualmente dispersas por todo el genoma, pero que no se agrupan en bloques.

a) Bloques dispersos de repeticiones en tándem: DNA minisatélite y DNA microsatélite

Se distinguen estos dos subgrupos en función del tamaño de la unidad de repetición. Constituyen conjuntamente entre el 5 y el 15% del genoma. Obviamente, a pesar de la similitud de nombre, no deben confundirse con el DNA satélite (pág. 113), del que se diferencian por estar localizados de forma dispersa por el genoma, aunque se asemejan a él en la presencia de repeticiones en tándem.

La importancia práctica del DNA minisatélite y microsatélite radica en su gran variabilidad entre individuos, que permite utilizarlo como marcador molecular en medicina forense, pruebas de paternidad y diagnóstico de enfermedades moleculares; estas aplicaciones se describen más adelante al hablar de polimorfismo genético (pág. 385).

a1) DNA minisatélite:

Está formado por repeticiones de 10 a 65 pb, ricas en G+C, agrupadas en tándem formando bloques relativamente grandes, de cientos o miles de repeticiones; esos bloques se encuentran repartidos por todo el genoma.

Algunas de estas repeticiones, denominadas "DNA minisatélite hipervariable", se caracterizan por su elevado polimorfismo (diferencias en secuencia y número de repeticiones), no sólo entre individuos de una especie, sino incluso entre los dos cromosomas homólogos de un individuo. Sus repeticiones, de 10 a 24 pb, presentan la secuencia consenso GGGCAGGANG (donde N representa cualquier nucleótido). Este DNA es importante por las aplicaciones derivadas de su polimorfismo, principalmente la obtención de **huellas de DNA**, donde una sonda para la secuencia consenso puede hibridar simultáneamente con múltiples loci minisatélites diferentes distribuidos por todos los cromosomas, dando lugar a un esquema o patrón de hibridación diferente para cada individuo.

a2) DNA microsatélite:

Se aplica este término cuando la unidad de repetición es inferior a 7 pb. Se presentan agrupadas en tándem en bloques de hasta 50 repeticiones, también distribuidos de forma dispersa por el genoma. Su polimorfismo se aplica igualmente para la obtención de huellas genéticas, para pruebas de identidad y estudios familiares.

Aunque las **secuencias teloméricas** se han clasificado anteriormente como DNA repetitivo agrupado (DNA satélite), por su ubicación en puntos concretos del cromosoma, algunos autores las consideran DNA microsatélite por el hecho de compartir con éste características de tamaño, número y disposición en tándem de las repeticiones.

b) Repeticiones dispersas: SINEs y LINEs

De nuevo, la longitud de la unidad de repetición sirve de criterio para subdividirlos:

b1) Secuencias SINE:

Las siglas proceden de *Short Interspersed Nuclear Elements*, **elementos nucleares dispersos cortos**. Se habla comúnmente de unidades de 100 a 500 pb, repetidas hasta 20 veces. Sin embargo, el ejemplo más característico, el elemento **Alu**, de 300 pb, aparece en el genoma humano entre 500.000 y un millón de veces. Su nombre se debe a que contiene un sitio diana para la enzima de restricción *Alu-I*. Se trata en realidad de una familia de secuencias, con un 85% de homología, que supone por sí sola alrededor de un 5% del genoma, de modo que, en promedio, existe una secuencia *Alu* cada 4 o 5 kb.

Se cree que los elementos *Alu* pueden haberse originado evolutivamente por la integración en el genoma de múltiples copias (*retroposones*) formadas por transcripción inversa del gen de RNA 7SL, un RNA citoplásmico pequeño (scRNA) componente de la *partícula de reconocimiento de señal* (pág. 324). A pesar de la homología (parcial) de secuencia, las secuencias *Alu* no ejercen la función de este scRNA.

b2) Secuencias LINE:

En este caso, se trata de **elementos nucleares dispersos largos** (*Long Interspersed Nuclear Elements*), con unidades de mayor tamaño, generalmente varios miles de pb, que se repiten hasta 50.000 veces en forma dispersa. La principal familia LINE en humanos la constituyen las secuencias **Kpn**, también llamadas **L1** o **LINE-1**. Se encuentran en el genoma de 50.000 a 100.000 repeticiones dispersas, de tamaños comprendidos entre 1,4 y 6,1 kb cada una, debido a su diversidad.

10.4 ESTUDIO EXPERIMENTAL DE LA COMPLEJIDAD EN EL GENOMA DE EUCARIOTAS

El descubrimiento y análisis del DNA repetitivo que, como se ha indicado, determina la "complejidad" del genoma eucariótico, se ha basado fundamentalmente en las **técnicas de reasociación** del DNA desnaturalizado. Esquemáticamente, este estudio implica las siguientes etapas:



Para obtener conclusiones al comparar la renaturalización de distintos tipos de DNA es conveniente que las moléculas que se reasocian sean de longitud regular. Esto se consigue (etapa 2) fragmentando el DNA genómico mediante sonicación de la disolución (ultrasonidos de alta intensidad) o haciéndola pasar a presión a través de un orificio pequeño, lo que ejerce fuerzas de cizalla que fragmentan las largas moléculas de DNA. Igualmente, conviene partir de la misma concentración de DNA, o en caso contrario debe compensarse su valor incluyéndolo en la variable *C_{ot}*, término definido a continuación. La reasociación se analiza (etapa 5) midiendo la fracción numérica de fragmentos que se han reasociado, en función del tiempo de templado. La proporción de DNA de doble hebra se determina espectrofotométricamente, o por otras técnicas que diferencien el DNA de hebra sencilla y doble.

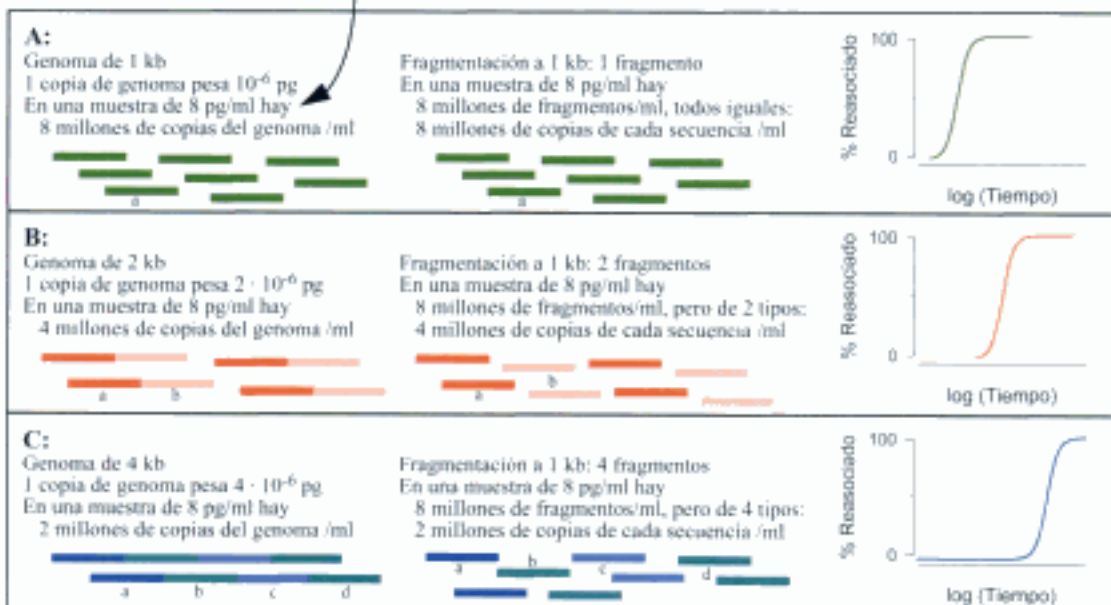
Como ejemplo de este tipo de estudio, se muestra la escisión de tres genomas hipotéticos sencillos (A, B y C), de tamaño diferente (1, 2 y 4 kb), en fragmentos de igual longitud (1 kb).

Ejemplo del cálculo, para el genoma "A":

El peso molecular de un par de bases es en promedio de 610, por lo que:

$$1 \text{ copia de genoma de } 1 \text{ kb} = 10^3 \text{ pb} = \frac{10^3}{6,023 \cdot 10^{23}} \text{ moles de pb} = \frac{10^3}{6,023 \cdot 10^{23}} \text{ mol} \cdot 610 \text{ g/mol} = 10^{-18} \text{ g} = 10^{-6} \text{ pg}$$

Entonces, si 1 copia de genoma pesa 10^{-6} pg, el número de copias/ml presentes en una muestra de 8 pg/ml es de $8 \cdot 10^6$



Cuanto mayor es el tamaño del genoma (A $\rightarrow$ B $\rightarrow$ C), menor es el n° de copias/ml tanto del genoma como de cada secuencia (a, b, c o d).

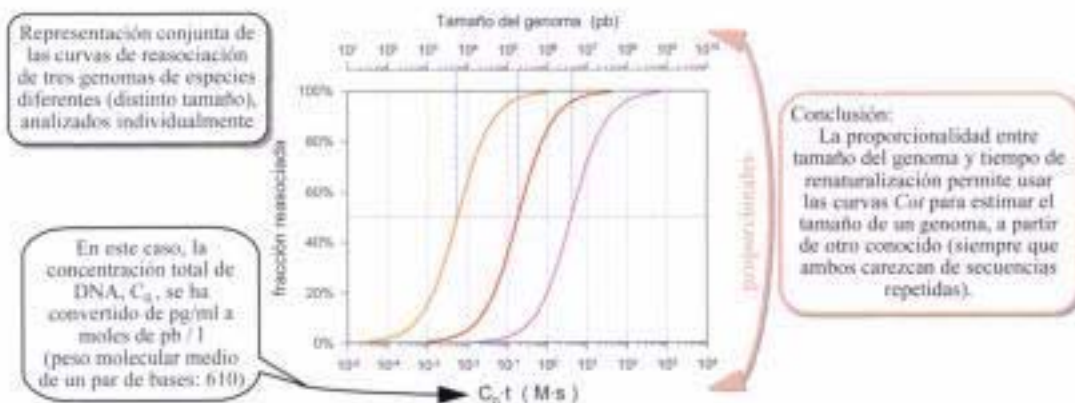
Obviamente, esto sólo es válido para una concentración de DNA fija, *C₀* (medida en pg/ml, en este caso 8 pg/ml).

En la renaturalización, cada fragmento tiene mayor probabilidad de interaccionar con su complementario cuanto mayor sea la concentración de ambos

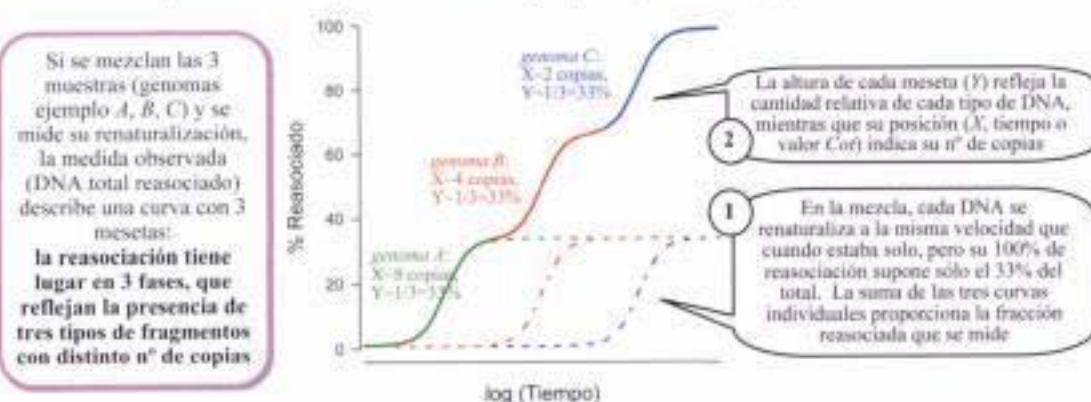
Conclusión: los genomas con mayor tamaño tardan más tiempo en renaturalizar

Para compensar el efecto que tendría la distinta concentración inicial de DNA en las muestras, se usa el producto $C_0 \times t$ en lugar del tiempo de reasociación, t . Este producto se conoce normalmente como "variable Cot ". La representación gráfica de la cinética de reasociación se hace, por tanto, con fracción o porcentaje reasociado en ordenadas y logaritmo de Cot en abscisas, y se habla coloquialmente de las "curvas Cot ".

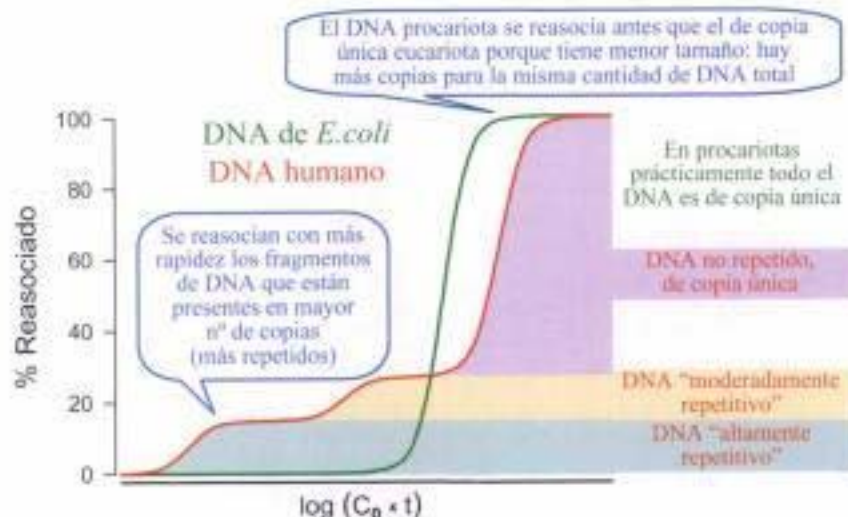
La observación de una correlación inversa entre la posición de la curva de reasociación y la magnitud del genoma de partida permite el empleo de las curvas Cot para obtener información sobre el tamaño de un genoma:



Además, las curvas Cot tienen especial interés porque aportan información acerca del contenido en DNA repetitivo de un genoma. Esto es debido a que, para una misma concentración de DNA total (o, en su defecto, corregida usando la variable Cot), los fragmentos de secuencias repetidas existen en mayor concentración, por lo que se reasocian más rápidamente. Se ilustra primero esta dependencia con un ejemplo simplificado, que parte de los anteriores:



De forma similar, cuando se analizan genomas reales las curvas Cot reflejan la presencia de secuencias con distinto número de repeticiones. Por ejemplo, para dos genomas típicos de un procariota (*E. coli*) y un eucariota (humano) se observa, de acuerdo con los criterios de interpretación explicados, que en el primero prácticamente todo el DNA es de copia única, mientras que el genoma humano posee un marcado componente repetitivo, que se agrupa bajo los términos "DNA altamente repetitivo" y "DNA moderadamente repetitivo".



Tema 11

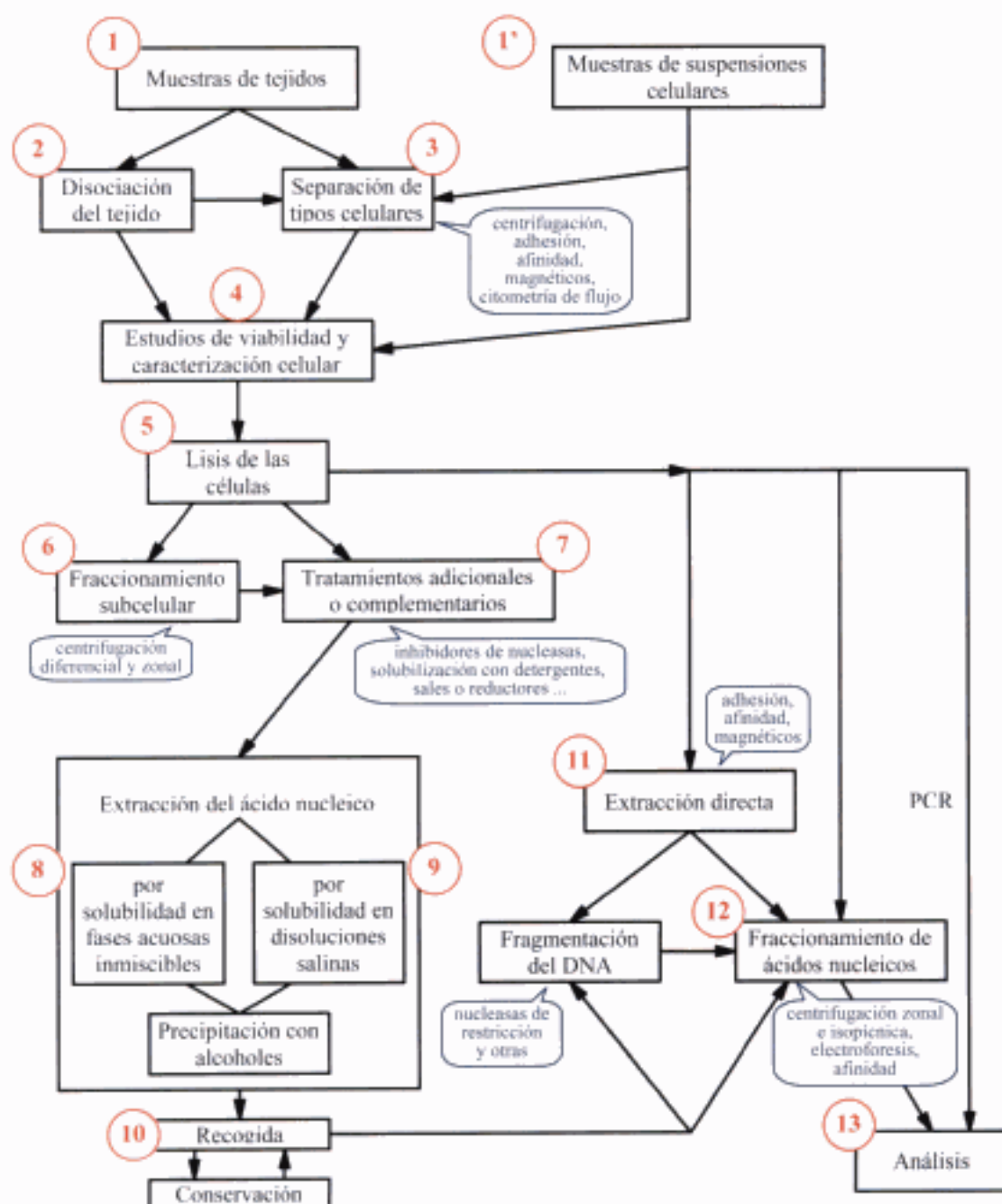
Preparación de muestras, extracción y análisis de ácidos nucleicos

11.1 Obtención y preparación preliminar de las muestras	119
11.1.1 Muestras de tejidos sólidos	119
11.1.2 Muestras de suspensiones celulares	119
11.1.2.1 Células embrionarias (diagnóstico pre-implantación)	119
11.1.2.2 Líquido o fluido amniótico	120
11.1.2.3 Células de tejidos o sangre fetales	120
11.1.2.4 Células de sangre o médula ósea adultas	120
11.1.2.5 Muestras bucales	120
11.1.2.6 Semen	120
11.1.2.7 Células en cultivo	121
11.1.3 Muestras celulares de líquidos biológicos extravasculares	121
11.1.3.1 Líquido cefalorraquídeo	121
11.1.3.2 Líquidos o derrames serosos	121
11.1.3.3 Líquido articular o sinovial	121
11.1.3.4 Muestras biológicas menos habituales	122
11.2 Disociación de la muestra tisular	122
11.3 Separación de células	122
11.3.1 Separación de células por métodos de sedimentación y centrifugación	122
11.3.1.1 Centrifugación diferencial	122
11.3.1.2 Centrifugación en gradiente de densidad	124
a) Métodos de barrera: centrifugación a través de un medio de densidad constante	124
b) Centrifugación zonal o de velocidad de sedimentación	124
c) Centrifugación isopícnica o de equilibrio de sedimentación	124
11.3.1.3 Elutriación centrífuga	125
11.3.2 Caracterización y separación de células por citofluorimetría de flujo	126
11.3.3 Separaciones directas de tipos celulares	127
11.3.3.1 Separaciones por afinidad con partículas magnéticas	128
11.4 Caracterización celular y medidas de viabilidad	129
11.4.1 Ensayos de exclusión	129
11.4.2 Recuento de células viables por incorporación de un marcador	130
11.4.3 Ensayos por medida de la función metabólica	131
11.5 Lisis de las células	132
11.6 Preparación de fracciones subcelulares	133
11.7 Tratamientos adicionales o complementarios	133
11.8 Extracción de ácidos nucleicos por solubilidad en fases inmiscibles	134
11.8.1 Extracción del DNA	134
11.8.2 Extracción del RNA	134
11.9 Purificación de ácidos nucleicos por precipitación salina diferencial	135
11.10 Recogida y conservación de las muestras	135
11.11 Procedimientos alternativos: extracción o análisis directos del DNA	135
11.12 Fraccionamiento de ácidos nucleicos	136
11.12.1 Ultracentrifugación	136
11.12.1.1 Centrifugación zonal	136
11.12.1.2 Centrifugación isopícnica	137
11.12.2 Electroforesis	137
11.12.2.1 Soporte para la electroforesis	138
11.12.2.2 Detección del DNA tras su separación ("revelado")	138
11.12.2.3 Calibración de la electroforesis para el cálculo del tamaño de los fragmentos de DNA	138
11.12.2.4 Electroforesis de campo pulsante	138
11.12.3 Métodos cromatográficos	139

Las técnicas derivadas de la Biología Molecular e Ingeniería Genética se utilizan de forma creciente tanto con fines básicos como aplicados, fundamentalmente al diagnóstico clínico, partiendo de muestras biológicas muy diversas. Ello requiere la extracción y análisis de RNA y DNA y, en su caso, proteínas u otras macromoléculas. Todos estos aspectos, como en todo proceso analítico, se desarrollan en tres fases:

- **Prealítica**, desde la obtención y preparación preliminar de la muestra inicial hasta su transporte al laboratorio.
- **Analítica** o de aplicación de procedimientos para la extracción de los ácidos nucleicos, sobre los que se realiza luego el análisis bioquímico-genético propiamente dicho: toma y tratamiento de datos, y recogida de resultados en un informe.
- **Postanalítica** o de validación técnica del informe analítico. Al conjunto de esta fase y la preanalítica se le llama "fase extra-analítica".

Estas fases, hoy día condicionadas por los continuos avances tecnológicos, instrumentales, de automatización e informáticos, han llegado a tal estado de desarrollo que pueden a veces enmascarar el carácter científico que subyace a todo el proceso analítico. Dentro del planteamiento de este libro, corresponde considerar solamente y de forma breve los aspectos esenciales de las fases preanalítica y analítica. Los detalles concretos, descritos en libros especializados, protocolos de laboratorio, etc., son objeto de estudio teórico-práctico durante los periodos de formación y de ejercicio profesional o científico en diversas áreas de laboratorio. Aunque de difícil delimitación y contenido (por ejemplo, a veces se llama *procesado de la muestra* a la preparación preliminar, pero otras se incluye en este término a la fase analítica), ambas fases se han integrado aquí, para su descripción, bajo un esquema unificado donde se intenta reunir todas las



posibilidades, aunque éstas puedan no darse simultáneamente en la realidad. Se establece así una serie de etapas más o menos consecutivas, conectadas entre sí, en las cuales se aplican técnicas muy diversas, en ocasiones sencillas y comunes, a veces muy complejas y específicas.

11.1 OBTENCIÓN Y PREPARACIÓN PRELIMINAR DE LAS MUESTRAS

La muestra celular se obtiene y prepara a partir de órganos, tejidos o fluidos biológicos, en función del objetivo perseguido (clínico o de identificación de individuos), tipo de célula (somática o germinal), tipo de genoma (nuclear o mitocondrial), tipo de estudio a realizar (secuencia del genoma, genes, productos génicos), etc. En cualquier caso, deben emplearse condiciones que *garanticen la calidad* de la muestra frente a las diversas variables que afectan a la fase preanalítica (recogida, anticoagulante empleado —en su caso—, manipulaciones para la conservación, transporte, etc.). En general, la preparación preliminar se hace rápidamente tras la obtención (dentro de 1 hora, como máximo 24 h), o la muestra se mantiene congelada hasta que llega al laboratorio.

11.1.1 Muestras de tejidos sólidos

Normalmente se obtienen mediante **biopsia** (método invasivo de recogida, bajo condiciones quirúrgicas, de células o porciones de tejido de un organismo vivo, con fines diagnósticos). Se liberan de restos de sangre, grasa o tejidos acompañantes por lavados con disolución salina fisiológica fría y se envían al laboratorio, según los casos, refrigeradas o congeladas, protegidas con hielo, nieve carbónica o nitrógeno líquido. La elección del tejido y tipo de extracción depende de muchos factores; entre otros, contenido en DNA, finalidad del estudio, o necesidad de amplificación por PCR de pequeñas muestras. En ocasiones se parte de tejidos fijados o de DNA antiguo.

Características generales de las principales muestras de tejidos y órganos de adultos	
Tejidos tumorales	Para analizar marcadores de membrana (glicoproteínas, receptores, etc.), de citosol (receptores hormonales, etc.) y de núcleo (oncogenes)
Tejido hepático	Rico en DNA (15 µg/mg); para enfermedades metabólicas
Tejido muscular	Pobre en DNA (3 µg/mg); para enzimopatías
Tejido intestinal	Para enzimopatías
Extensiones de piel	Pobre en DNA (3 µg/mg). Se obtiene de zonas no vellosas
Raiz de pelo	Cantidad mínima de DNA (0,5 µg/raiz); con fines forenses; difícil de estudiar
Huesos	Pulpa de los dientes o médula de otros huesos. El tejido óseo se tritura por métodos físicos
Tejidos fijados	Muestras anatomopatológicas, fijadas con formaldehído (formol)
Tejidos incluidos en parafina	Se debe eliminar la parafina empleando disolventes, por ejemplo, xileno o tolueno
DNA antiguo (restos humanos y animales)	Interés arqueológico y antropológico. Debe evitarse la contaminación por DNA exógeno e inhibidores de Taq-polimerasa (que alteran el análisis por PCR)
DNA atrapado	Desde hace millones de años. La resina vegetal (ámbar) mantiene el DNA en perfecto estado

11.1.2 Muestras de suspensiones celulares

11.1.2.1 Células embrionarias (diagnóstico pre-implantación)

El óvulo fertilizado origina en el útero por sucesivas mitosis estructuras de 2, 4 y 8 células totipotentes (8 *blastómeros*, que forman la *mórula*) y luego de 16, 32 y 64 células pluripotentes; esta última estructura se llama *blastocisto* o *blastodermo*, y a partir de ella se inicia el desarrollo. Para observar si el embrión es portador de un trastorno genético se puede obtener DNA extrayendo 1 ó 2 *blastómeros* (lo que no afecta al desarrollo de las otras células de la *mórula*).

El *blastodermo* pasa a formar una capa externa de células (*trofoblasto*) y una masa celular interna (*blastocelo* o *saco vitelino primitivo*). A unos 5 días de la fecundación se produce la implantación, cuando el *trofoblasto* invade la pared del endometrio (que reviste internamente el útero), dando lugar a la maduración de las *vellosidades coriónicas*, que penetran en la placenta para permitir la circulación fetal. Las células de un polo del *blastocelo* originan los sacos embrionario y vitelino, mientras que el resto del *blastocelo* forma el embrión. De la masa celular que constituye este último se pueden obtener células madre pluripotentes, de elevado interés potencial para trasplantes, pues pueden dar lugar en cultivo a células especializadas de distintos tipos.

11.1.2.2 Líquido o fluido amniótico

A partir de la 1ª semana de gestación, el embrión queda protegido al estar inmerso en el *líquido amniótico*, contenido dentro de una membrana llamada saco amniótico o *amnios*. La composición del líquido amniótico, cuyo volumen llega hasta 1.100 ml a término, depende de un equilibrio dinámico de intercambios con el feto y la madre, y contiene células epiteliales que aportan unos 65 ng DNA/ml. La obtención de una muestra, mediante punción abdominal a través de la pared del útero, se denomina *amniocentesis*. Ésta se usa a partir de las 14 semanas de gestación para el diagnóstico prenatal de anomalías cromosómicas, enfermedades metabólicas congénitas, defectos del desarrollo (inmadurez pulmonar del feto), estudio del sexo, etc. Las muestras del 3º trimestre sirven para valorar la gravedad de la eritroblastosis fetal y predecir el síndrome de dificultad respiratoria del neonato.

11.1.2.3 Células de tejidos o sangre fetales

Como alternativa al líquido amniótico se utilizan muestras de **vellosidades coriónicas**, tejido fetal procedente del *corion* (membrana más externa del embrión, que contacta con la placenta). Éstas constituyen la mejor fuente de DNA fetal. Se obtienen por biopsia a las 8-12 semanas de gestación y han de ser cultivadas.

Se pueden obtener también muestras de **sangre fetal**, dentro del seno materno, empleando dos procedimientos: La *fetoscopia* (método similar a la endoscopia, válido tanto para exploración como para toma de muestras) se utiliza entre las 16 y 21 semanas de gestación, para determinar hemoglobinopatías, coagulopatías, análisis de cariotipo, etc. Una técnica alternativa más sencilla, la *cordocentesis* o *funiculocentesis*, permite obtener sangre fetal directamente del cordón umbilical al 3º trimestre de embarazo; sus eritrocitos (nucleados) y linfocitos se utilizan para diagnóstico prenatal sin interferencias con el feto.

11.1.2.4 Células de sangre o médula ósea adultas

Las muestras de **sangre periférica** son las más empleadas, comúnmente de sangre venosa (unos 40 µg DNA/ml) recogida sobre anticoagulante. Se utiliza la fracción celular total o sus distintos tipos, una vez separados por alguno de los métodos disponibles, en especial la centrifugación según sus diferencias de tamaño o densidad.

Se utilizan también **manchas de sangre**, que se obtienen de sangre fisiológica (talón, dedo...) secada sobre un papel o cartón, o bien de origen accidental (prendas, vehículos, utensilios...). La mancha se raspa y extrae por lavado con disolución salina. Se utilizan para metabolopatías, identificación, fines forenses, etc. Se pueden destacar las *tarjetas Guthrie*, para diagnóstico neonatal, y las tarjetas de *papel FTA*, material especialmente tratado en el que se seca la sangre, se lisan las células y se extraen sus componentes, quedando fijado exclusivamente el DNA. Estas tarjetas poseen la ventaja de su fácil transporte y conservación.

Las **muestras de médula ósea roja** (activa en hematopoyesis) se obtienen, bajo anestesia local, por aspiración mediante *punción medular*, generalmente en el esternón o en la cadera, usándose para observaciones citológicas al microscopio y estudios genéticos e histológicos.

11.1.2.5 Muestras bucales

Los **raspados y enjuagues bucales**, obtenidos por métodos no invasivos, con cepillo suave o algodón y disolución salina, contienen células epiteliales de descamación con suficiente DNA genómico para su detección, previa amplificación por PCR.

La **saliva** es de uso creciente por su facilidad de obtención (toma directa o presencia sobre la piel, sellos, colilla de cigarrillo, etc.). Las células epiteliales que incorpora, procedentes de descamación, pueden aportar cantidad suficiente de DNA (8-9 µg en una toma de muestra con hisopo) para su análisis con fines forenses o estudio de la paternidad.

11.1.2.6 Semen

Contiene millones de espermatozoides (5-7% del volumen total, con 1,3 pg DNA por célula) formados en los testículos, en suspensión en una fracción líquida de secreciones de las vesículas seminales (50% en volumen), la próstata (20% en volumen), las glándulas bulbouretrales y el epidídimo. Se recoge de varias maneras, por ejemplo: (a) en preservativo, para estudios de funcionamiento del tracto genital masculino, marcadores de fertilidad, etc., (b) por aspiración directa de la vagina, para investigación criminal por agresión sexual, o (c) por lavado salino de manchas secas en ropa, piel, pelo, etc. Las muestras frescas se deben analizar antes de 1 h de su recogida, manteniéndolas a temperatura corporal durante su transporte. Los espermatozoides se pueden separar del resto de células (que incluyen las de la víctima, en casos de asalto sexual) mediante lisis diferencial de éstas con SDS (dodecilsulfato sódico) y proteinasa K.

11.1.2.7 Células en cultivo

Las células aisladas o procedentes de tejidos u órganos disgregados se pueden cultivar en medios especiales. El cultivo celular constituye *per se* un área de estudio independiente, con requerimientos especializados en cuanto a equipo, procedimientos, sustratos, nutrientes y factores de crecimiento, etc. Pueden citarse como ejemplos líneas celulares clónicas (genética y metabólicamente uniformes con la célula inicial) y cultivos de células de tejidos, con un potencial de aplicación como alternativa al trasplante de órganos y tejidos o para la reparación de lesiones tisulares; por ejemplo, fibroblastos y queratinocitos para preparar "parches" de piel artificial (dermis y epidermis, respectivamente), hepatocitos para patologías hepáticas terminales, y células de miocardio para lesiones cardíacas.

11.1.3 Muestras celulares de líquidos biológicos extravasculares

Se incluyen en este apartado algunas muestras con interés diagnóstico por sí mismas o complementario de las anteriores.

11.1.3.1 Líquido cefalorraquídeo

Es un fluido resultante de la ultrafiltración del plasma (de ahí su analogía en composición) y que recoge los productos de desecho del metabolismo cerebral. Se encuentra en dos compartimentos que recubren el cerebro y la médula espinal, comunicados entre sí: el *sistema ventricular* (30 ml) y el *espacio subaracnoideo*, delimitado por las membranas piamadre y aracnoides (120 ml). Se produce y reabsorbe continuamente (con un recambio de ~500 ml/día) en los ventrículos de los dos hemisferios cerebrales. Sirve como soporte protector del sistema nervioso central, para que "flote" el cerebro evitando su apoyo directo sobre la superficie ósea. Se puede obtener una muestra por *punción lumbar*. Aunque su aspecto normal es transparente, en condiciones patológicas aumenta su turbidez por la presencia de células (esencialmente leucocitos, eritrocitos, microorganismos, etc.), en función del tipo y gravedad de la enfermedad. Se utiliza para detectar enfermedades del sistema nervioso central, así como para detectar marcadores tumorales.

11.1.3.2 Líquidos o derrames serosos

Se localizan en las cavidades del cuerpo que rodean los órganos, delimitadas por dos membranas, la *parietal* (externa) y la *visceral* (interna, superficie del órgano). Estas membranas están formadas por tejido conjuntivo y una rica red de capilares sanguíneos y linfáticos, a través de cuya pared se filtra el plasma cuando se altera el balance hidrostático u osmótico; de ahí la analogía de composición entre plasma y líquidos serosos. En la práctica clínica, se distinguen dos tipos de ultrafiltrado: *transudados*, de origen no inflamatorio, con pocas células (menos de 300/ml), y *exudados*, debidos a derrames inflamatorios y con muchas células (más de 1.000/ml). Las muestras, normalmente transparentes, sólo se toman en situaciones anormales, presentando entonces un grado de turbidez variable en función de su contenido en leucocitos y eritrocitos, fundamentalmente.

	volumen normal	ubicación	obtención por punción o paracentesis
Líquido peritoneal o ascítico	100 ml	cavidad abdominal	punción peritoneal
Líquido pleural	10-30 ml	cavidad pleural (se acumula en la base de los pulmones)	<i>toracocentesis</i> o <i>pleurocentesis</i> (punción de la pared torácica o del espacio pleural)
	Contiene células de muchos tipos y es de interés clínico en enfermedades inflamatorias, infecciosas, marcadores tumorales, etc.		
Líquido pericárdico	20-50 ml	espacio pericárdico	<i>pericardiocentesis</i>
	Se usa con menor frecuencia, ya que presenta más complicaciones que los anteriores		

11.1.3.3 Líquido articular o sinovial

Es un ultrafiltrado del plasma, viscoso, con un alto contenido en proteínas y ácido hialurónico, localizado en la cavidad articular para conferir un carácter *lubricante* frente a la fricción propia de las articulaciones durante el movimiento. Con un volumen normal de 1 a 4 ml, su contenido celular tiene interés tanto en procesos mecánicos como en procesos inflamatorios patológicos articulares. Se obtiene por *artrocentesis*, es decir, aspiración tras punción en el espacio articular.

11.1.3.4 Muestras biológicas menos habituales

Pueden contener células y, por tanto, DNA, una serie de muestras de interés en casos concretos: humor acuoso (cámara anterior del ojo), humor vítreo (detrás del iris), lágrimas, endolinfa (tubo espiral del oído interno) y perilinfa (canales del equilibrio), moco cervical, líquidos folicular, oviductal y vaginal, heces, secreciones nasales, uñas, etc.

La **orina**, que normalmente no contiene células, se utiliza para medida de metabolitos en enfermedades hereditarias relacionadas con el metabolismo de los aminoácidos (fenilcetonuria, alcaptonuria, cetonuria de aminoácidos ramificados, cistinuria, etc.), de las bases purínicas (gota) y pirimidínicas (aciduria orótica), del grupo hemo (porfirias), etc.

11.2 DISOCIACIÓN DE LA MUESTRA TISULAR

En los tejidos, las células se encuentran atrapadas en una organización tridimensional, con interacciones o adhesiones tanto entre ellas como con la matriz extracelular (gel hidratado formado por fibras proteicas, como colágeno, y cadenas glucídicas complejas). El mejor rendimiento en cuanto a disociación de células viables se alcanza a partir de tejidos fetales o neonatales. Pueden señalarse **tres estrategias**:

Fragmentación mecánica: Mediante corte, troceado o triturado de la muestra original en disolución amortiguadora de pH, para obtener células individuales o fragmentos tisulares pequeños que faciliten su obtención.

Disgregación química: La rotura de la matriz extracelular y de las uniones intercelulares se suele realizar por homogeneización isotónica o por sonicación (exposición a ultrasonidos, ondas sonoras o acústicas no percibidas por el oído humano por poseer una frecuencia superior a 20 kHz), generalmente en presencia de detergentes (iónicos, como el dodecilsulfato sódico o SDS, o no iónicos, como Tween-20 y Triton X-100). La eliminación de cationes divalentes, en especial el Ca^{2+} (con agentes quelantes como EDTA), también desestabiliza la adhesión de células entre sí y con la matriz extracelular.

Digestión enzimática: Las proteínas de la matriz se digieren con proteasas como collagenasa, proteinasa K, pepsina o tripsina. Dependiendo del tipo de muestra, esta digestión puede requerir largos tiempos de incubación (por ejemplo, células en cultivo o raíces de pelo).

11.3 SEPARACIÓN DE CÉLULAS

La obtención de muestras **celulares** adecuadas para la fase analítica requiere variados procedimientos de separación, fraccionamiento y caracterización, basados normalmente en las propiedades de cada tipo celular existente en la muestra. Su elección en distintas etapas de las fases preanalítica y analítica depende de muchos factores: muestra de partida, grado de separación deseado, conservación de la viabilidad, operaciones y técnicas a aplicar posteriormente, etc. Además de los métodos más comunes, que se exponen a continuación, existen otros procedimientos específicos, pero que no han alcanzado un uso general (por ejemplo, reparto en sistemas bifásicos acuosos, electroforesis de flujo libre, etc.).

11.3.1 Separación de células por métodos de sedimentación y centrifugación

La separación celular puede hacerse por sedimentación a *gravedad unidad*, bajo la fuerza de la gravedad terrestre, pero lo habitual es acelerar el proceso aumentando muchas veces esta fuerza mediante la centrifugación. Al realizarse ésta en el seno de un medio líquido, la separación no sólo depende de la fuerza centrífuga, sino de otros factores que modifican la sedimentación y que dependen de las características de la célula (o partícula en general) y del medio. El conjunto define la **velocidad de sedimentación** de la célula (pág. 123)

11.3.1.1 Centrifugación diferencial

Normalmente, el planteamiento más sencillo es la separación por centrifugación, comenzando con valores bajos de *g* (*precentrifugación*). En la práctica, sólo se pueden separar así células que muestren grandes diferencias de tamaño, al menos de 10 veces (por ejemplo, plaquetas, mucho más pequeñas que las restantes células sanguíneas). Por ejemplo, en el caso de sangre con anticoagulante, a 200g los eritrocitos sedimentan en la zona inferior, los leucocitos aparecen en la interfase y queda como sobrenadante un plasma rico en plaquetas (*PRP*). Se puede preparar plasma pobre en plaquetas (*PPP*) empleando mayor velocidad (alrededor de 3.000g).

1 Recuerdese:

La separación por centrifugación depende de la masa celular (m), del radio de giro del tubo, o radio del rotor, donde se coloca la muestra (r) y de la velocidad alcanzada por la centrifuga (ω o v):

$$F_c = m \cdot a = \frac{m \cdot v^2}{r} = \frac{m \cdot e^2}{r \cdot t^2} = \frac{m \cdot \phi^2 \cdot r^2}{r \cdot t^2} = \frac{m \cdot \phi^2 \cdot r}{t^2} = m \cdot \omega^2 \cdot r = \\ = m \cdot (2\pi)^2 \cdot v^2 \cdot r = 39,48 \cdot m \cdot v^2 \cdot r$$

Siendo:

F_c = fuerza centrífuga

m = masa de la célula

$a = v^2/r$ = aceleración; si es a gravedad unidad, $a = g = 9,8 \text{ m/s}^2 = 980 \text{ cm/s}^2$

$v = e/t$ = velocidad lineal de la muestra en el tubo (cm/min)

r = radio de giro, medido entre el punto medio del tubo de centrifuga y el eje de rotación (cm)

e = espacio lineal recorrido, si el movimiento fuese lineal (cm)

t = tiempo de sedimentación empleado al centrifugar (min)

ϕ = ángulo recorrido en el movimiento de rotación (radianes = ángulo cuyo arco es igual a r); ($\phi = e/r$)

$\omega = \phi/t$ = velocidad angular, pues el movimiento es de rotación (radianes/min = min^{-1})

v = velocidad de giro (revoluciones por segundo = rps, o revoluciones por minuto = rpm); ($\omega = 2\pi v$)

La fuerza centrífuga se suele expresar en forma de múltiplos de la fuerza de gravedad, llamándose entonces **fuerza centrífuga relativa (FCR)**. (Coloquialmente, se denomina número de "ges" y se indica como 100g, 1.000g, etc. o bien 100xg, 1.000xg.)

$$FCR = \frac{F_c}{F_g} = \frac{a}{g} = \frac{39,48 \cdot m \cdot v^2 \cdot r}{m \cdot 980 \text{ cm/s}^2} = 0,0403 \cdot v^2 \cdot r \text{ cm}^{-1} \cdot \text{s}^2$$

Si, como ocurre normalmente, v se mide en revoluciones/min (rpm) y r en cm, se pueden expresar como

$$v = \left(\frac{v}{\text{rpm}} \right) \cdot \text{rpm} = \left(\frac{v}{\text{rpm}} \right) \cdot \frac{1}{\text{min}} = \left(\frac{v}{\text{rpm}} \right) \cdot \frac{1}{60 \text{ s}} \quad r = \left(\frac{r}{\text{cm}} \right) \cdot \text{cm}$$

y en consecuencia, el valor anterior de FCR viene dado por

$$FCR = 0,0403 \cdot \left[\left(\frac{v}{\text{rpm}} \right)^2 \cdot \frac{1}{3600 \text{ s}^2} \right] \cdot \left[\left(\frac{r}{\text{cm}} \right) \cdot \text{cm} \right] \cdot \text{cm}^{-1} \cdot \text{s}^2 = 11,19 \cdot 10^{-4} \cdot \left(\frac{v}{\text{rpm}} \right)^2 \cdot \left(\frac{r}{\text{cm}} \right)$$

$$(\text{Habitualmente escrito como } FCR = 11,19 \cdot 10^{-4} \cdot v^2 \cdot r = \frac{v^2 \cdot r}{89365})$$

Ejercicio resuelto: El nº de g alcanzado en un rotor de centrifuga, de 5 cm de radio, que gira a una velocidad de 10.000 rpm es: $FCR = (10.000)^2 \cdot 5 / 89.365 = 5.600g$

2 La velocidad a la que sedimenta una célula (u otra partícula) en un medio líquido viene definida por la Ley de Stokes:

$$\text{Velocidad de sedimentación} = 4r^2 (\rho_c - \rho_m) FCR / 18 \eta$$

con las variables medidas en: cm/s $(\text{cm})^2$ g/cm^3 poise (1 poise = $1 \text{ g} \cdot \text{s/cm}^2$)

Es decir, es proporcional a la fuerza centrífuga generada (**FCR**), al tamaño de la célula (su radio, r) y a la diferencia entre la densidades de la célula (ρ_c) y la del medio de suspensión (ρ_m), y es inversamente proporcional a la viscosidad del medio (η).

Las propiedades de sedimentación de una célula o partícula se indican a menudo mediante su *coeficiente de sedimentación*, s :

$$\text{Coeficiente de sedimentación} = s = \frac{\text{velocidad de sedimentación}}{\text{aceleración centrífuga}}$$

expresado en unidades Svedberg (1 S = 10^{-13} segundos).

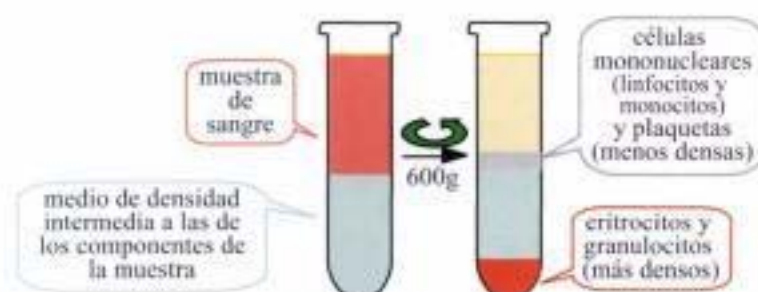
En algunos casos se centrifuga en un medio con el que previamente se ha formado un gradiente de densidad creciente hacia la parte inferior del tubo. En este caso, la célula sedimenta hasta que la densidad del medio es igual a la suya propia; en ese punto, la velocidad de sedimentación se hace cero y la célula no sedimenta más, es decir, ha alcanzado el equilibrio.

11.3.1.2 Centrifugación en gradiente de densidad

a) Métodos de barrera: centrifugación a través de un medio de densidad constante

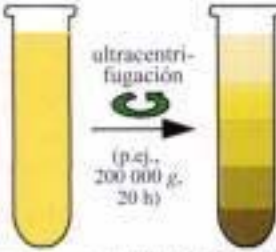
Para conseguir separaciones más eficaces (por ejemplo, la de los leucocitos respecto al resto de las células sanguíneas) se centrifuga la muestra sobre un lecho de densidad intermedia entre la de los tipos celulares que se quieren separar. Se emplean para ello medios comerciales como Ficoll-Paque, Lymphoprep y otros muchos, formados generalmente por mezclas de Ficoll (un polisacárido sintético) y metrizamida (un compuesto sintético yodado).

Se dispone de varios medios con densidades adecuadas para la separación de tipos celulares concretos; por ej., Nycoprep-1.077 para células mononucleares, Nycoprep-1.068 para monocitos, Polymorphoprep para células polimorfonucleares, o Nycoprep-1.063 para plaquetas.



b) Centrifugación zonal o de velocidad de sedimentación

Para separar tipos celulares cuyo coeficiente de sedimentación difiere poco (p.ej., linfocitos y monocitos) es preciso acudir al uso de medios que forman un gradiente de densidad, lo que favorece la concentración de cada tipo celular en una banda o zona estrecha; por ello, a esta técnica se la denomina *centrifugación zonal*. El gradiente de densidad se genera, antes de añadir la muestra (gradiente *preformado*), mediante varias disoluciones de distinta concentración de un compuesto adecuado (Ficoll, sacarosa, albúmina, suero fetal bovino...). La preparación puede hacerse de forma manual (dando lugar a un gradiente discontinuo) o empleando un aparato *formador de gradientes* (gradiente continuo).

Gradiente preformado		Gradiente autoformado															
Gradiente discontinuo	Gradiente continuo	Gradiente continuo															
Ejemplo de preparación, mediante mezcla de varias disoluciones (válido para centrifugación zonal e isopícnica, dependiendo de las densidades elegidas y del tiempo de centrifugación)  <table><tr><td>Se depositan capas sucesivas de distinta concentración de sacarosa (y distinta densidad)</td><td>10%</td><td>1.038</td></tr><tr><td></td><td>15%</td><td>1.059</td></tr><tr><td></td><td>20%</td><td>1.081</td></tr><tr><td></td><td>25%</td><td>1.104</td></tr><tr><td></td><td>30%</td><td>1.127</td></tr></table> (% peso/vol) (g/cm³)	Se depositan capas sucesivas de distinta concentración de sacarosa (y distinta densidad)	10%	1.038		15%	1.059		20%	1.081		25%	1.104		30%	1.127	Ejemplo de preparación, mediante mezcla de 2 disoluciones en un dispositivo <i>formador de gradientes</i>  Se llena el tubo con la mezcla gradual de dos disoluciones de distinta concentración (y densidad)	Ejemplo de preparación, mediante centrifugación. El gradiente se va formando a la vez que se separan los componentes de la muestra. (El ejemplo, CsCl, no se usa para células, sino para separar ácidos nucleicos) Se ultracentrifuga una disolución concentrada de CsCl que sedimenta gradualmente formando un gradiente de concentración y de densidad  ultracentrifugación (p.ej., 200 000 g, 20 h) $d = 1,70 \text{ g/cm}^3$ arriba: $1,66 \text{ g/cm}^3$ abajo: $1,76 \text{ g/cm}^3$
Se depositan capas sucesivas de distinta concentración de sacarosa (y distinta densidad)	10%	1.038															
	15%	1.059															
	20%	1.081															
	25%	1.104															
	30%	1.127															

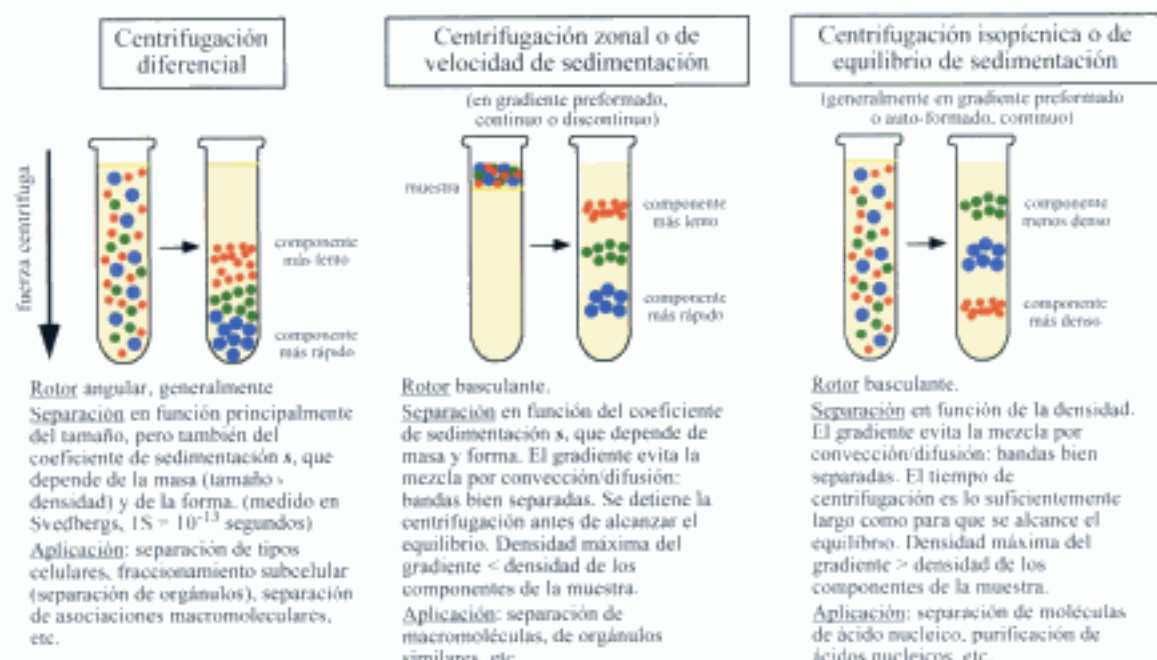
Un pequeño volumen de muestra se deposita sobre el gradiente y se centrifuga a alta velocidad durante un tiempo corto. Habitualmente se emplea un gradiente cuya densidad máxima es inferior a la de las células, con lo que éstas podrían alcanzar el fondo del tubo si se centrifugase durante tiempo suficiente. Se emplean siempre tiempos menores, por lo que el resultado es dependiente del tiempo de centrifugación; dicho de otro modo, no se alcanza el equilibrio de sedimentación. Con esta técnica se consigue, por ejemplo, separar todos los tipos celulares sanguíneos, purificar espermatozoides viables, separar células viables y no viables de tejidos desagregados y muestras con células en suspensión, etc.

c) Centrifugación isopícnica o de equilibrio de sedimentación

Se pueden separar células en función exclusivamente de su densidad empleando otra variante de centrifugación, llamada *isopícnica* (palabra de origen griego que significa *de igual densidad*). Se realiza también sobre un gradiente de densidad, pero en este caso el tiempo de centrifugación es lo suficientemente largo como para que se alcance el equilibrio de sedimentación (entre la fuerza centrífuga, el empuje hidrostático de la célula y su difusión). Para

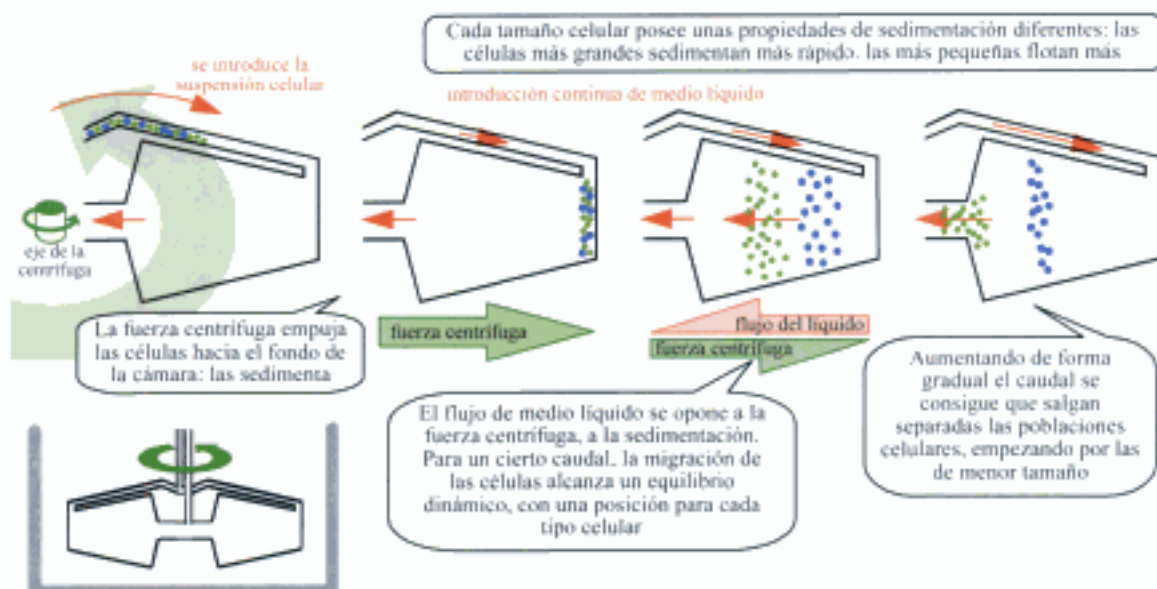
conseguirlo, se usan gradientes continuos que cubren todo el intervalo de densidades de las células; en el fondo del tubo la densidad del medio ha de ser mayor que la de las células más densas. De esta forma, independientemente del tiempo de centrifugación, las células nunca sedimentarán en el fondo, sino que alcanzan una posición estable intermedia en el gradiente.

Con una variante peculiar de esta técnica se consigue purificar linfocitos de su mezcla con monocitos, aprovechando la capacidad de estos últimos para fagocitar microesferas de oro coloidal, con lo que su densidad se hace artificialmente muy superior y sedimentan más abajo del tubo.



11.3.1.3 Elutriación centrífuga

La **elutriación** consiste en la separación de células (o partículas en general) de acuerdo con su velocidad de sedimentación. El proceso original, consistente en la realización de ciclos sucesivos de sedimentación y decantación, fue facilitado por la incorporación al sistema de un flujo del propio medio líquido de suspensión, en sentido contrario a la sedimentación (contra la gravedad). Este fue finalmente combinado con el efecto de la fuerza centrífuga, dando lugar a la modalidad de elutriación de mayor uso actual, la **elutriación centrífuga**. Aquí, las células se ven sometidas a dos fuerzas opuestas: por una parte, la fuerza centrífuga y, por otra, la fuerza de arrastre por el flujo continuo del medio en sentido contrario.



El resultado es un método versátil, rápido, eficaz y relativamente suave para separar subpoblaciones celulares según su tamaño a partir de una mezcla en suspensión. Admite números grandes de células e, incluso, permite obtener

una mayor proporción de células viables respecto a la muestra original (al separarse las células dañadas, de distinta densidad). Se aplica para obtener gran diversidad de subpoblaciones enriquecidas, a partir de poblaciones celulares hepáticas, células en proliferación (sincronizadas en la fases M, G₁, S o G₂), células sanguíneas y otros tipos celulares.

11.3.2 Caracterización y separación de células por citofluorimetría de flujo

Es hoy día un método de separación importante en investigación biomédica y en el laboratorio clínico, especialmente por su capacidad para el análisis automático de suspensiones celulares. El instrumento empleado se basa en el transporte de una suspensión celular (células sanguíneas, aspirado medular, tejidos disociados, etc.), impulsada por un flujo de disolución isotónica, al punto de medición o cámara de flujo. En su aplicación habitual, la citometría de flujo analiza y separa células previamente marcadas con un fluorocromo.

Recuérdese:

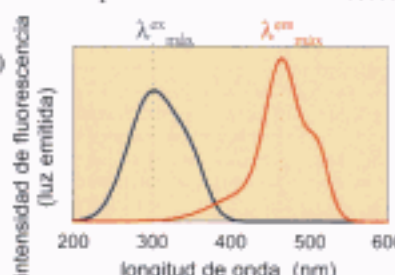
Fluorimetría: fundamento

Ciertas moléculas poseen la propiedad de ser fluorescentes, es decir, capacidad de absorber energía de una radiación electromagnética (luz visible o ultravioleta), que las hace pasar a un estado electrónico excitado desde el cual vuelven luego al estado fundamental, liberando energía también en forma de luz. Como la energía de la radiación electromagnética emitida es siempre inferior a la absorbida, la longitud de onda de la luz emitida es superior a la de la luz absorbida:

$$\lambda^{em} > \lambda^{ex}$$

(por *emisión* y *excitación*, respectivamente)

Cada compuesto con esta propiedad, denominado *fluoróforo*, posee espectros característicos de excitación y de emisión (el de excitación es normalmente similar al de absorción). Si además es un colorante, se denomina *fluorocromo*.



Espectro de excitación: se mide la intensidad de luz emitida al ir variando la longitud de onda con la que se ilumina la muestra (λ^{ex}).
Espectro de emisión: se mide la intensidad de luz emitida a cada longitud de onda (λ^{em}), manteniendo fija la longitud de onda con la que se ilumina.

El empleo de fluoróforos y fluorocromos como moléculas marcadoras es de una gran utilidad, pues ofrece la posibilidad de detectar distintas moléculas y células en muy pequeña cantidad y sin los riesgos asociados al marcaje con isótopos radiactivos. En concreto, en Biología Molecular se utilizan principalmente dos tipos de fluorocromos: agentes intercalantes que se unen al ácido nucleico y sirven así para detectarlo en geles de electroforesis o gradientes de densidad (tales como bromuro de etidio, naranja de acridina y colorantes Hoechst), y fluorocromos unidos de forma covalente a sondas de hibridación y a cebadores de PCR (fluoresceína, rojo Texas, rodaminas, etc.).

La detección se puede realizar mediante:

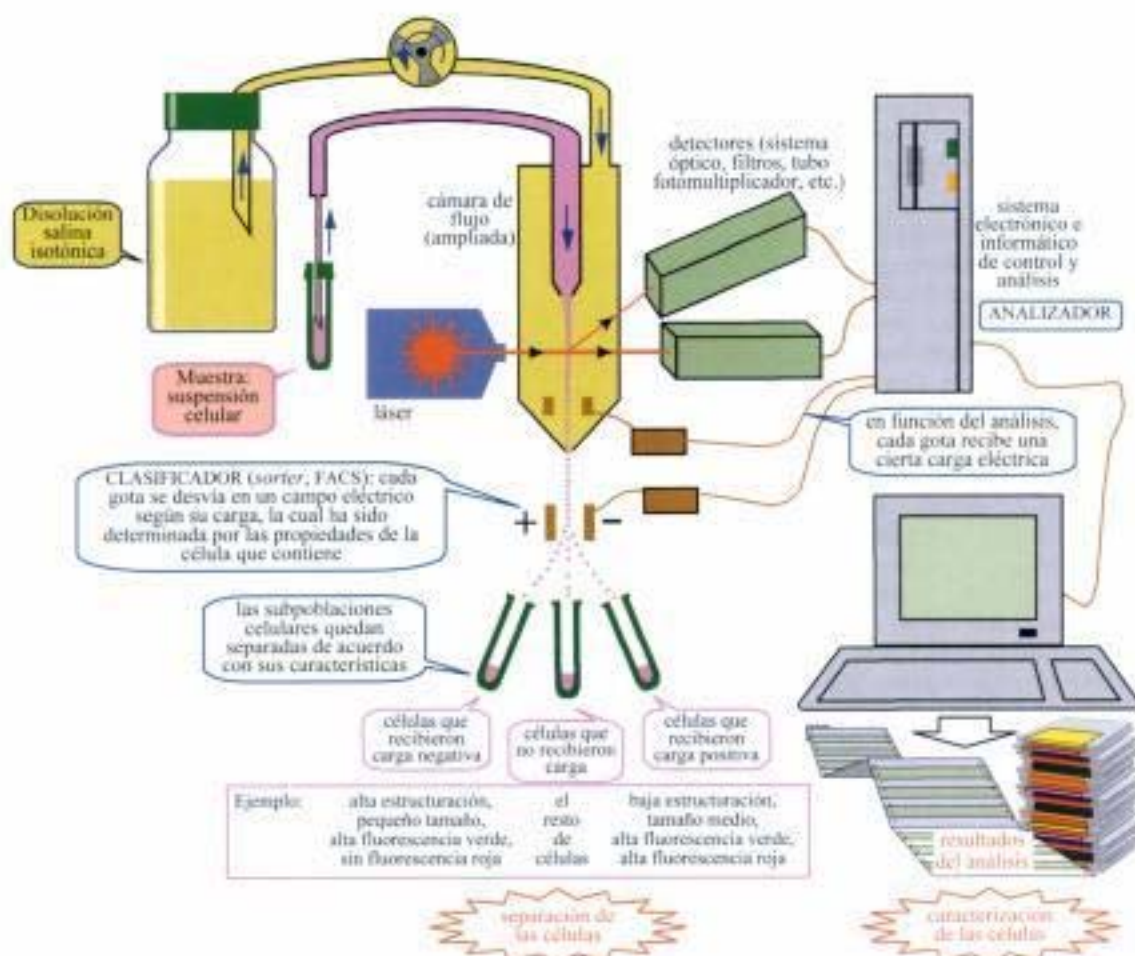
- espectrofluorímetro y lector de microplacas: miden la fluorescencia promedio de la muestra
- densitómetro de fluorescencia: en geles de electroforesis, *blots* o cromatografías en placa
- microscopio de fluorescencia: permite observar secciones de tejido o células individuales, de forma semicuantitativa
- citómetro de flujo: permite medir células individuales, de forma cuantitativa

El procedimiento de citometría de flujo consta de dos etapas: caracterización y separación. De hecho, algunos instrumentos (analizadores) sólo realizan la primera, mientras que otros (clasificadores) son capaces además de llevar a cabo la segunda.

a) Caracterización de la población celular: analizador, **citómetro de flujo** o **citofluorímetro**

El sistema consigue que la suspensión celular llegue en condiciones de flujo laminar, formando un chorro muy fino que, como consecuencia de la dilución, contiene las células individuales en sucesión. Éstas pasan así de una en una a través de un haz luminoso láser, cuya longitud de onda permite la excitación de los marcadores fluorescentes incorporados previamente a la célula. En la luz emergente de cada célula se analiza la dispersión y la intensidad de la fluorescencia (emisión).

Parámetro medido	Propiedades celulares analizadas
Dispersión frontal de la luz (refracción) o FSC	Tamaño (diámetro y razón de volumen núcleo/citoplasma)
Dispersión perpendicular de la luz (reflexión) o SSC	Arquitectura interna (grado de estructuración subcelular, contenido de orgánulos, granularidad, rugosidad ...)
Intensidad de fluorescencia (a una o varias λ)	Cantidad de fluorocromos marcadores previamente incorporados a la célula. Comúnmente se usan anticuerpos, marcados con fluorocromos, contra antígenos de la superficie celular o bien colorantes o sondas que se fijan al DNA



b) Separación de células: clasificador de células activado por fluorescencia (FACS)

La corriente de líquido, tras pasar por el punto de detección, se transforma en una sucesión de gotas muy pequeñas (generalmente, como consecuencia de una vibración aplicada por ultrasonidos a la cámara de flujo), las cuales contienen como máximo una sola célula. Al mismo tiempo, en función de los resultados del análisis anterior de la célula, el clasificador aplica automáticamente a cada gota un pulso de voltaje (entre ± 50 y ± 150 V) que le proporciona una carga eléctrica.

El usuario define previamente el pulso que se ha de aplicar a cada gota (y, por tanto, el signo y cuantía de la carga que adquirirá), en función de criterios sobre los parámetros del análisis en el citómetro. El sistema informático que controla analizador y clasificador permite tal decisión en el breve intervalo transcurrido entre el paso de cada célula por el detector y la formación de la gota que la contiene. Al mismo tiempo, todos los datos de cada célula se almacenan para su estudio posterior.

A continuación, las gotas cargadas pasan por un campo eléctrico, donde se desvían de acuerdo con su carga para caer en distintos tubos de recolección de muestra; de esta manera se consigue la separación de subpoblaciones celulares de acuerdo con los criterios definidos (combinación de tamaño, estructuración y fluorescencia).

Como meros ejemplos de las innumerables aplicaciones de esta técnica se pueden citar: a) El estudio de la superficie celular con anticuerpos monoclonales, siendo posible seleccionar una célula, con un antígeno específico de superficie, de entre 1.000 que no lo tengan y clasificar miles de células por segundo. b) La valoración del contenido en DNA y RNA de una célula (p.ej., en estudios del ciclo celular). c) La determinación de su forma y tamaño, por ejemplo, en neoplasias.

11.3.3 Separaciones directas de tipos celulares

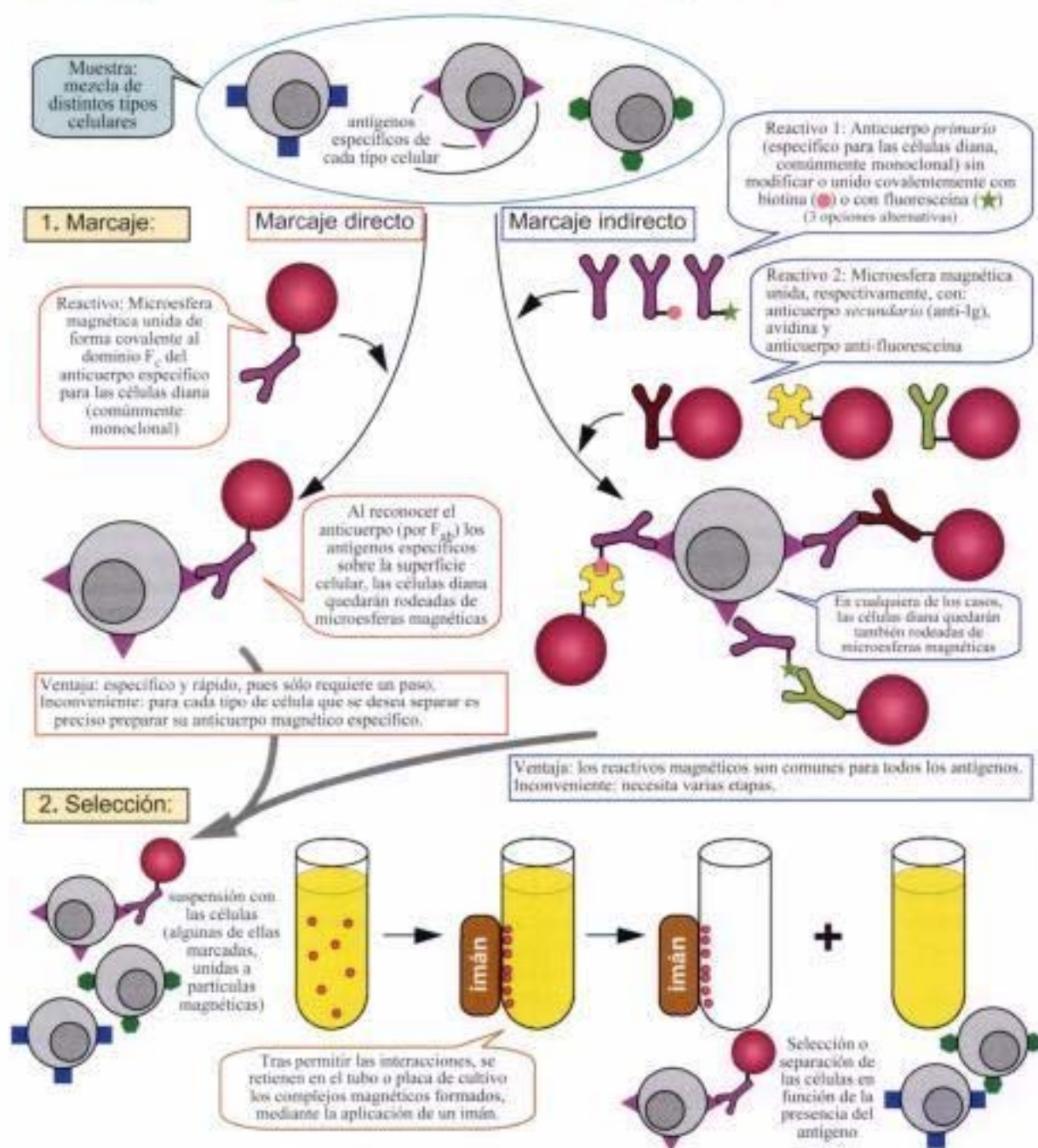
Aun siendo un caso particular, es de uso habitual y no puede dejar de mencionarse la posibilidad de aislar ciertos tipos celulares gracias a su tendencia a adherirse fuertemente a la superficie de un soporte de vidrio o plástico (**adhesión celular**), generalmente la propia placa de cultivo.

Con la llegada de los **anticuerpos monoclonales** este principio se pudo ampliar a prácticamente cualquier tipo de célula. La posibilidad de disponer de estos anticuerpos ha contribuido notablemente a la separación de células en el laboratorio clínico y de investigación. Su exquisita selectividad en la unión a epítopos antigénicos de la superficie ha permitido diseñar métodos diversos para aislar y caracterizar de forma directa los tipos celulares presentes en una mezcla heterogénea. Estos métodos pueden aplicarse, por ejemplo, para definir las células a trasplantar en caso de enfermedades graves (células precursoras de médula ósea evitando células T maduras inmunocompetentes), para

determinar el fenotipo de tejidos para trasplantes, establecer los tipos celulares responsables de procesos en la respuesta inmune, cuantificar subpoblaciones celulares, etc.

El uso de anticuerpos ha dado lugar a una gran diversidad de **métodos inmunoanalíticos** y de separación, que se pueden considerar parte de los **métodos de afinidad**. Por ejemplo, las células de un tipo pueden unirse, a través de sus antígenos, a anticuerpos monoclonales previamente fijados en superficies de plástico (tubos, placas de cultivo, varillas, microesferas ...) o matrices de colágeno, polisacárido, etc., lo que permite la separación de ese tipo celular. Como ejemplo singular, de entre la multitud de variantes desarrolladas, se comentan a continuación los métodos basados en el uso de microesferas magnéticas.

11.3.3.1 Separaciones por afinidad con partículas magnéticas



Selección positiva: Las microesferas magnéticas se unen a las células que interesa conservar para su estudio. Es, por su simplicidad, el abordaje recomendado. Sin embargo, requiere disponer de un método para romper la unión entre la célula y la microesfera, anticuerpo o avidina (dependiendo del uso posterior que se quiera dar a las células).

Selección negativa: Se marcan con las microesferas todos los tipos celulares excepto el de interés. Es la única solución cuando no se dispone de anticuerpos específicos o existe interferencia de los anticuerpos o las microesferas con posteriores experimentos. Posee la ventaja de que se obtienen las células intactas, sin nada unido a ellas, y así se pueden utilizar directamente.

A menudo, el mejor resultado se obtiene con varias selecciones negativas previas, que enriquecen la muestra antes de proceder a la selección positiva final.

Estos procedimientos, de gran sencillez y por ello en expansión actualmente, hacen uso de **microesferas** disponibles comercialmente en varios tamaños (entre 0,5 y 10 μm , es decir, menores que las células) y con distintas posibilidades de aplicación. Están formadas por un material superparamagnético (óxido de hierro) recubierto de una capa fina de un polímero plástico que permite la adsorción o la unión covalente de distintas moléculas. Frecuentemente se les une un anticuerpo, para formar *anticuerpos magnéticos* o *microesferas inmuno-magnéticas*, o bien avidina o estreptavidina, proteínas ambas con gran afinidad por la biotina (*microesferas magnéticas de afinidad*).

Más adelante se exponen aplicaciones similares de estas microesferas magnéticas para la extracción directa de ácidos nucleicos.

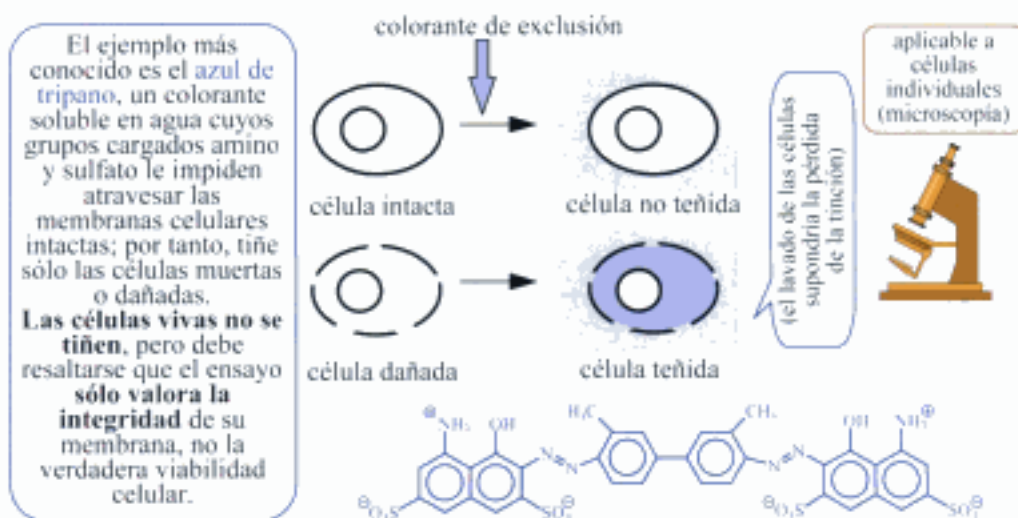
11.4 CARACTERIZACIÓN CELULAR Y MEDIDAS DE VIABILIDAD

Es importante proceder en esta fase preanalítica a la caracterización de las células por alguno de los procedimientos disponibles en la biología moderna: **microscopía óptica**, como punto de partida más obvio; **microscopía de contraste de fase**, en paralelo al anterior, para observar con más detalle la superficie celular; **microscopía electrónica de barrido** (SEM, de *scanning electron microscopy*), para definir la superficie celular y la naturaleza de las interacciones célula-célula, o **microscopía electrónica de transmisión** (TEM, de *transmission electron microscopy*), para observar cortes celulares y examinar los orgánulos o inclusiones. También se caracterizan las células por **citometría de flujo** o el uso de **microesferas magnéticas**, técnicas ya comentadas.

Es igualmente importante confirmar el estado vital de las muestras celulares durante una u otra etapa de la fase preanalítica. El concepto de **viabilidad** alude a la capacidad de una célula para realizar sus procesos fisiológicos y bioquímicos, especialmente en lo que respecta a su metabolismo y su capacidad de división. Sin embargo, en la práctica el término es relativo, pues se utiliza con diversos criterios; por ejemplo, es común que se hable de viabilidad refiriéndose simplemente a la integridad de la célula, o bien a su actividad metabólica o, finalmente, a su capacidad de división (proliferación). Para todos ellos existen numerosos **ensayos de viabilidad**; los más empleados son:

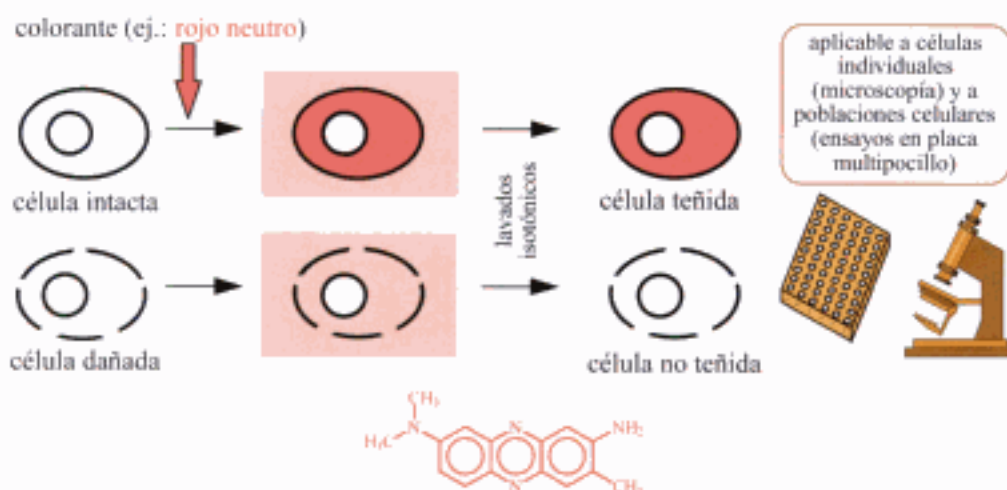
11.4.1 Ensayos de exclusión

Corresponde al recuento de células que no incorporan un colorante.

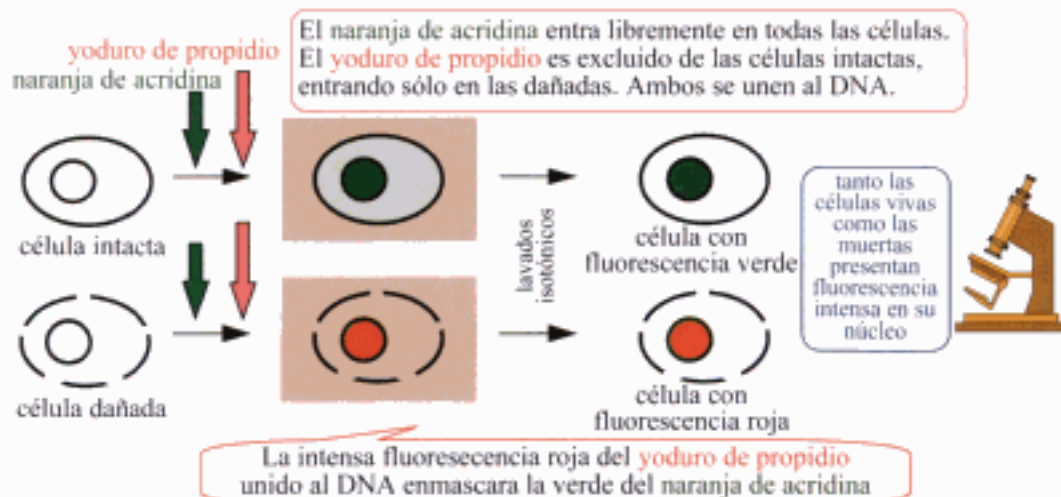


11.4.2 Recuento de células viables por incorporación de un marcador

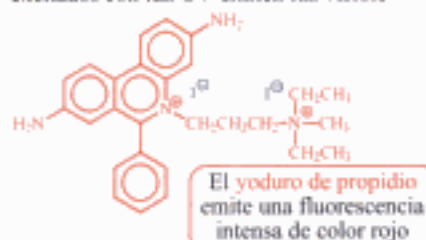
Método alternativo, basado en la entrada de un compuesto a través de la membrana de células intactas.



En algunos casos, el colorante (generalmente un fluorocromo) se une a los ácidos nucleicos, en especial al DNA.



Ambos compuestos son agentes intercalantes: su estructura aromática plana favorece su inserción entre los pares de bases apilados del DNA. Poseen propiedades fluorescentes: al ser excitados con luz UV emiten luz visible

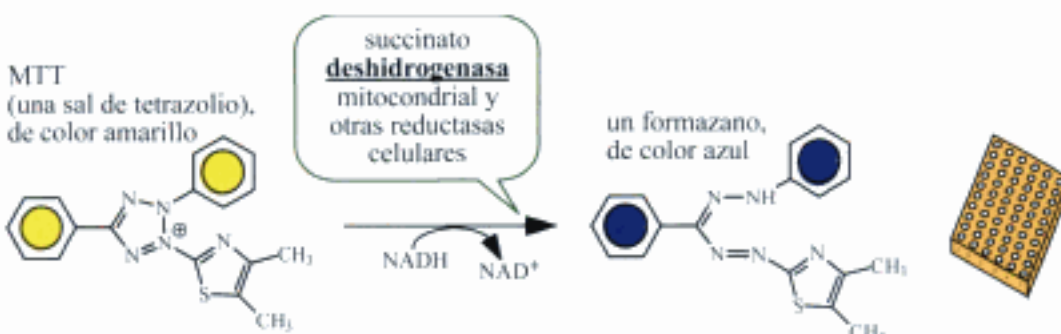
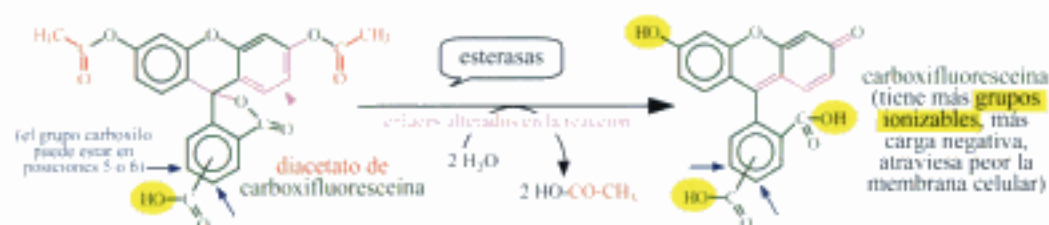
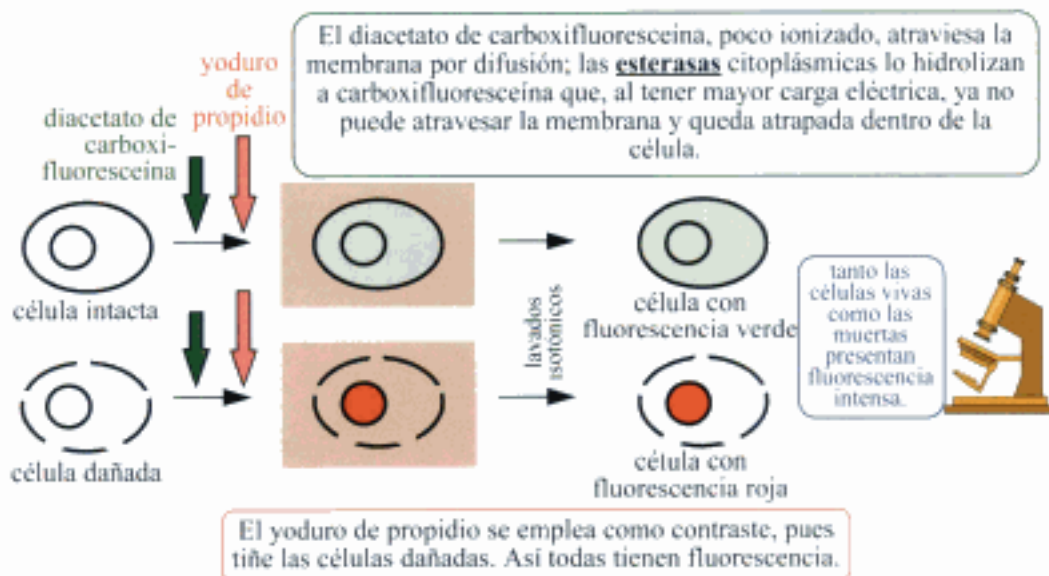


El naranja de acridina emite una fluorescencia de distinto color según se encuentre unido a ácidos nucleicos de hebra doble o sencilla:

- verde con DNA de doble hebra
- tonos intermedios (de amarillo a naranja) con DNA parcialmente desnaturizado
- roja con DNA de hebra sencilla (desnaturizado completamente) o con RNA

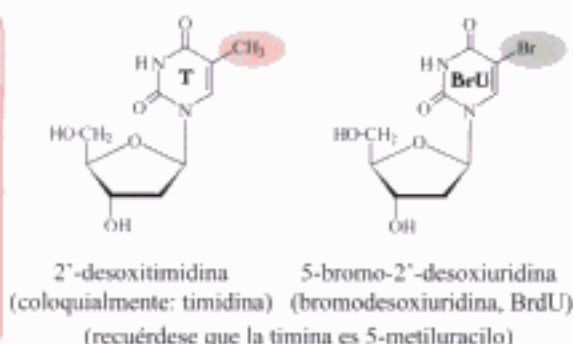
11.4.3 Ensayos por medida de la función metabólica

- **Actividades enzimáticas:** La presencia en las células de actividades enzimáticas genéricas, tales como esterasas y deshidrogenasas, permite plantear ensayos de viabilidad basados en la generación de productos con grupos cromóforos o fluoróforos, procedentes de la ruptura enzimática de sustratos sintéticos. Una alternativa, para tipos celulares concretos, es la determinación de una actividad enzimática específica, en cuyo caso es especialmente importante evitar que las condiciones experimentales afecten a dicha enzima.



- **Síntesis proteica:** incorporación de radiactividad a las proteínas formadas durante el cultivo de las células en presencia de aminoácidos marcados (generalmente, con ^{14}C , ^3H o $[^{35}\text{S}]\text{Met}$).
- **Síntesis de DNA:** incorporación de nucleósidos marcados. Refleja la viabilidad en términos de capacidad de crecimiento o, más estrictamente, proliferación celular.

Si se añade desoxitimidina marcada con ^3H (timidina tritiada) al medio de cultivo, ésta se incorpora al DNA durante la replicación. Se detecta la radiactividad de las células, debida a su DNA



BrdU es un análogo estructural de dT, pues el átomo de Br tiene un tamaño similar al del grupo metilo. Por ello se incorpora al DNA nuevo durante la replicación, en lugar de dT. Se detecta mediante anticuerpos específicos, generalmente marcados con un fluorocromo, en métodos de inmunofluorescencia o de citometría de flujo

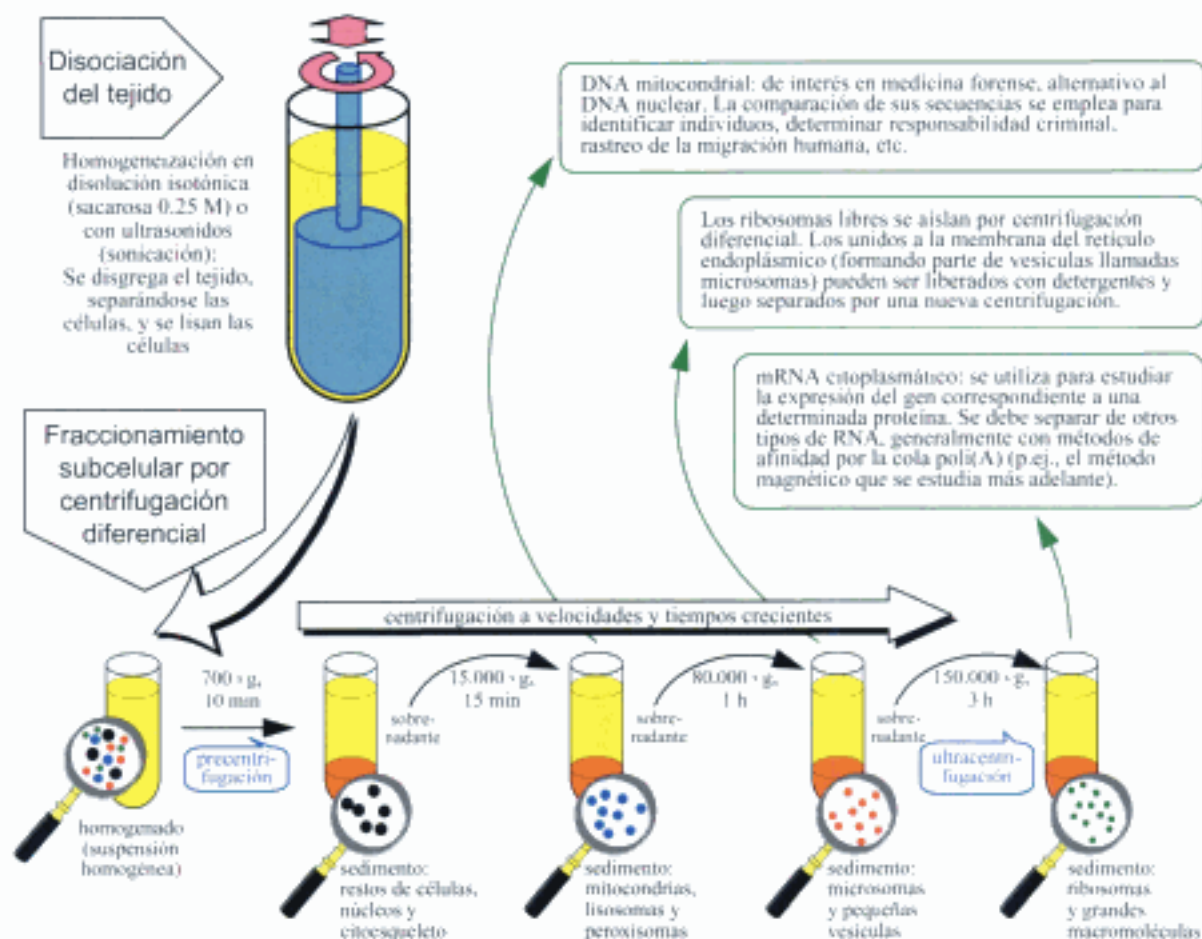
11.5 LISIS DE LAS CÉLULAS

Esta etapa supone el punto de enlace entre la preparación de la muestra celular y la de **muestras de genoma o ácido nucleico**; comienza, por tanto, la fase analítica, que comprende el fraccionamiento subcelular, la extracción directa o el fraccionamiento de ácidos nucleicos y la fragmentación del DNA.

En general, la lisis celular se realiza con un medio hipotónico; en algunos casos esta etapa es simultánea con la preparación de la muestra (p.ej., homogeneización). La lisis se puede aplicar tanto a células en suspensión como a las fijadas por adhesión a placas de cultivo. En la preparación de células sanguíneas es frecuente la aplicación de reactivos de lisis diferencial (sólo de determinados tipos celulares). Así, para el recuento y estudio de leucocitos se lisan los eritrocitos con un medio ácido (por ejemplo, líquido de Türk) o una disolución de cloruro amónico.

11.6 PREPARACIÓN DE FRACCIONES SUBCELULARES

El procedimiento general, empleado clásicamente, parte del sobrenadante que resulta de la disociación mecánica de un tejido por homogeneización en un medio isotónico y posterior lisis celular. Mediante centrifugación diferencial se separan sucesivamente núcleos, orgánulos y otras estructuras celulares.



11.7 TRATAMIENTOS ADICIONALES O COMPLEMENTARIOS

Es esencial combinar las operaciones precedentes y posteriores con una serie de tratamientos adicionales o complementarios. El lisado celular se puede someter a tratamientos que prevengan su destrucción o eliminen moléculas aún presentes, etc. Para mantener la integridad de las muestras, especialmente de algunos tejidos (por ejemplo, páncreas y bazo), se ha de evitar la acción de nucleasas endógenas (DNasas y RNasas); de ahí la conveniencia de usar muestras frescas, mantenidas congeladas o añadir inhibidores de nucleasas. Para su solubilización se acude a la rotura celular por homogeneización o resuspensión en presencia de detergentes, ya mencionados, y/o al empleo de disoluciones de alta fuerza iónica. Estos tratamientos son especialmente importantes para extraer el RNA, lo que también requiere de forma específica la presencia de agentes disociantes (como el tiocianato de guanidina, inhibidor de RNasas) o reductores (por ejemplo, 2-mercaptoetanol).

Para la preparación de RNA todas las disoluciones empleadas deben tratarse previamente con dietilpirocarbonato (DEPC), que inactiva covalentemente las RNasas; igualmente, se incluyen durante la preparación de la muestra y purificación del RNA otros inhibidores de RNasas. Debe también ponerse especial atención a la eliminación de cualquier resto de DNA, generalmente añadiendo una DNasa.

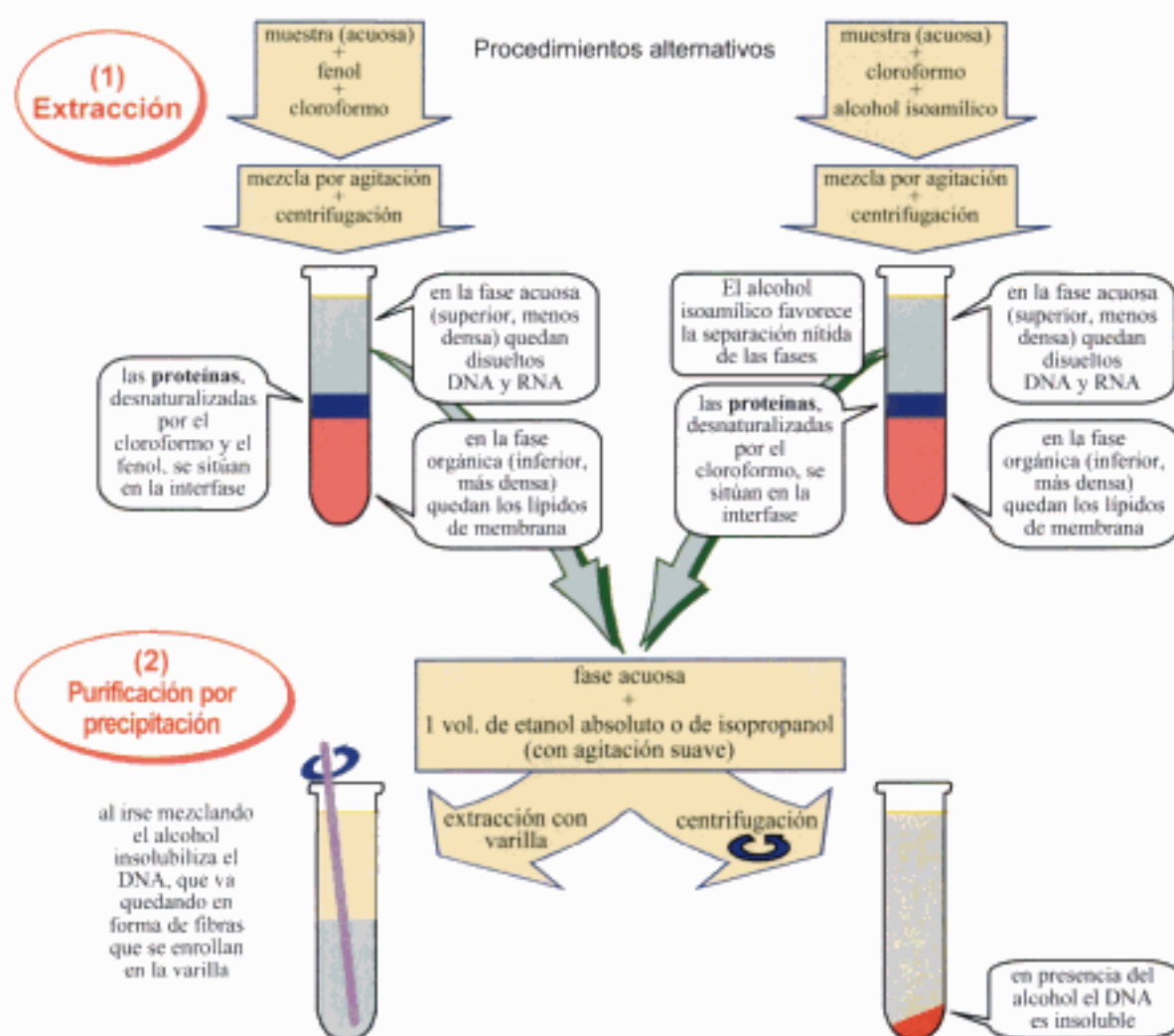
Igualmente, tanto en la obtención de RNA como de DNA, el material de vidrio y plástico debe estar completamente libre de ácidos nucleicos y nucleasas, y debe evidentemente evitarse la contaminación por el analista mediante el uso de guantes y precauciones como el manejo separado de las muestras de distinto origen.

11.8 EXTRACCIÓN DE ÁCIDOS NUCLEICOS POR SOLUBILIDAD EN FASES INMISCIBLES

En general, dada su complejidad macromolecular, los ácidos nucleicos son difíciles de separar de los componentes de membrana, proteína residual o añadida (tales como las enzimas de restricción) y otras moléculas solubles (por ejemplo, polisacáridos). Los métodos drásticos alteran sus estructuras y sus características. De ahí que sean muchos los métodos suaves elaborados para eliminar la impureza de las preparaciones, todos ellos basados en las propiedades físico-químicas del ácido nucleico. Para la eliminación de proteínas asociadas se utilizan enzimas proteolíticas, como la proteinasa K.

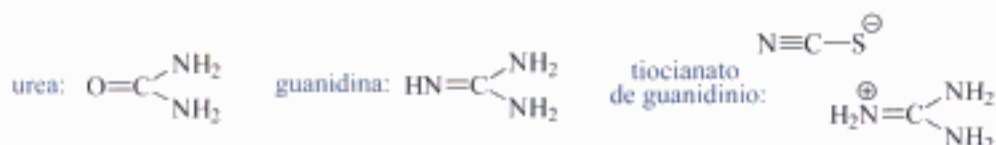
11.8.1 Extracción del DNA

El protocolo básico consiste en el tratamiento del homogenado con disolventes orgánicos (fenol, cloroformo o ambos).



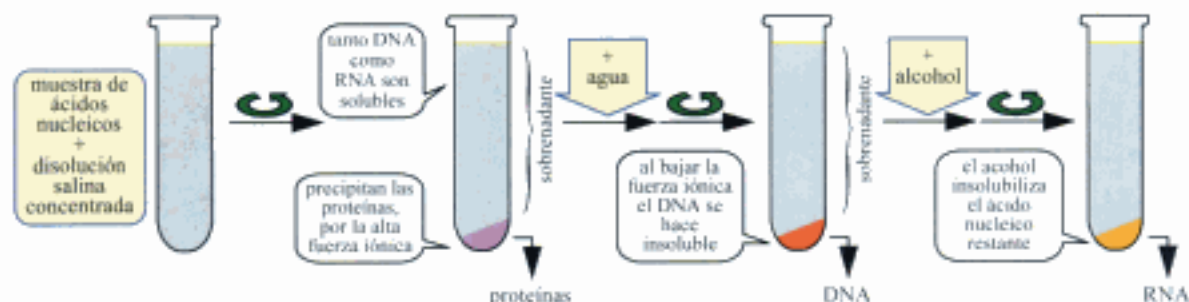
11.8.2 Extracción del RNA

La extracción del RNA también se realiza a partir del homogenado, en presencia de tiocianato de guanidinio, por adición de fenol, cloroformo y acetato sódico de pH=5, seguido de agitación vigorosa y centrifugación; el RNA aparece en la fase acuosa superior. A partir de ella se obtiene el RNA por precipitación con etanol. Existen otras técnicas más simples que no requieren dicha extracción, como el empleo de soluciones de urea y cloruro de litio, que también inactivan RNasas y solubilizan el RNA.



11.9 PURIFICACIÓN DE ÁCIDOS NUCLEICOS POR PRECIPITACIÓN SALINA DIFERENCIAL

Se trata de un procedimiento alternativo al anterior, basado en las diferencias de solubilidad de DNA y RNA en función de la concentración de sales. Tanto DNA como RNA son solubles en disolución salina concentrada, mientras que a baja concentración de sales sólo es soluble el RNA.



Para esta precipitación salina se suele utilizar NaCl, pero puede también emplearse acetato sódico saturado, en combinación con alcoholes orgánicos, conduciendo a resultados incluso mejores que con fenol y cloroformo, sobre todo para cantidades pequeñas de DNA.

11.10 RECOGIDA Y CONSERVACIÓN DE LAS MUESTRAS

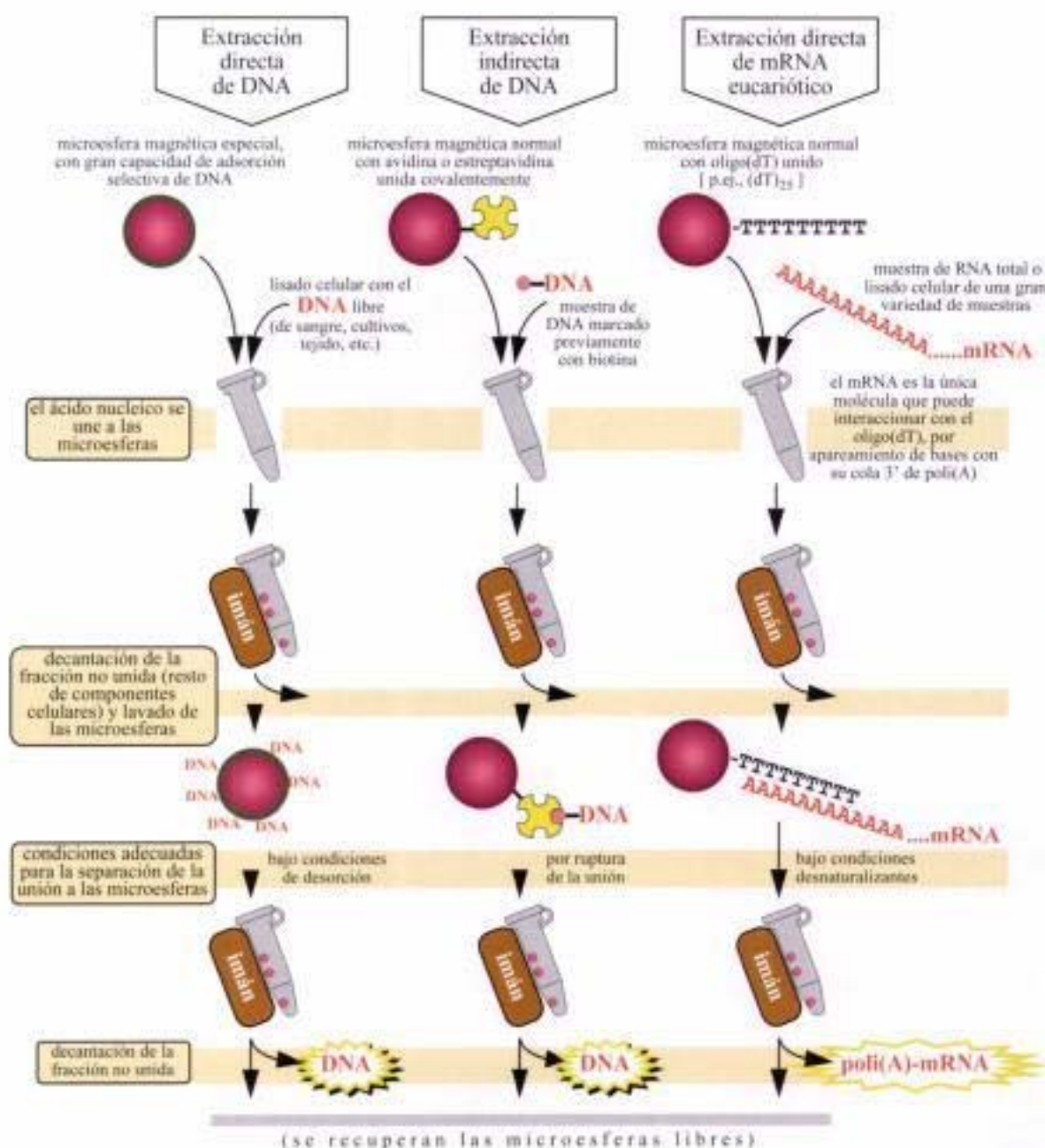
Una vez separados y extraídos, los ácidos nucleicos deben ser conservados adecuadamente. El precipitado seco de ácidos nucleicos o el "ovillo" de fibras de DNA obtenido con la varilla pueden disolverse en disolución amortiguadora (alcalina, en el caso del DNA) y conservarse a 4 °C, o disolverse en agua y mantenerse congelados para posteriores manipulaciones enzimáticas. El DNA genómico en disolución no se debe congelar, para evitar la rotura de las moléculas por las fuerzas de cizalla provocadas al solidificarse el hielo y al descongelarse. Por el contrario, el RNA sí puede conservarse congelado a -70 °C en disolución acuosa, cuidando de añadir un inhibidor de RNasas.

11.11 PROCEDIMIENTOS ALTERNATIVOS: EXTRACCIÓN O ANÁLISIS DIRECTOS DEL DNA

A veces, no es posible utilizar los procedimientos descritos, dado el bajo contenido en DNA de algunas muestras, pero, sin embargo, puede estudiarse el DNA de la muestra sin necesidad de manipulación alguna, salvo la lisis de las células. Este es el caso de la amplificación y detección del DNA mediante la técnica de PCR (ver tema posterior). Por otro lado, cada vez es más frecuente y eficaz la *extracción directa* del DNA con métodos más específicos, fáciles, rápidos de realizar y fiables. No se requiere la manipulación previa de la muestra, con lo que el DNA y RNA se pueden extraer de forma directa a partir de muestras heterogéneas y de pequeño volumen. Existen varios procedimientos.

Para muestras forenses se ha descrito, con resultados prometedores, el empleo de un intercambiador iónico (Chelex) y una suspensión de partículas de sílice como **materiales de adsorción** adecuados para la extracción directa de DNA, en especial mitocondrial. También se extiende el uso para pruebas de identidad de las *tarjetas FTA*, papel especialmente tratado que permite absorber una muestra (gota de sangre), lisar sus células y digerir sus proteínas, quedando el DNA intacto fijado en el papel y así se puede conservar indefinidamente la muestra.

También se emplean cada vez más **microesferas magnéticas**, análogas a las de separación celular, sobre cuya superficie se adsorben directamente los ácidos nucleicos o se unen de forma covalente otras moléculas con afinidad por ellos. Se consiguen así métodos simples de extracción que se pueden aplicar a una amplia variedad de muestras (tejidos animales o vegetales, células cultivadas, suspensiones celulares, etc.).



11.12 FRACCIONAMIENTO DE ÁCIDOS NUCLEICOS

11.12.1 Ultracentrifugación

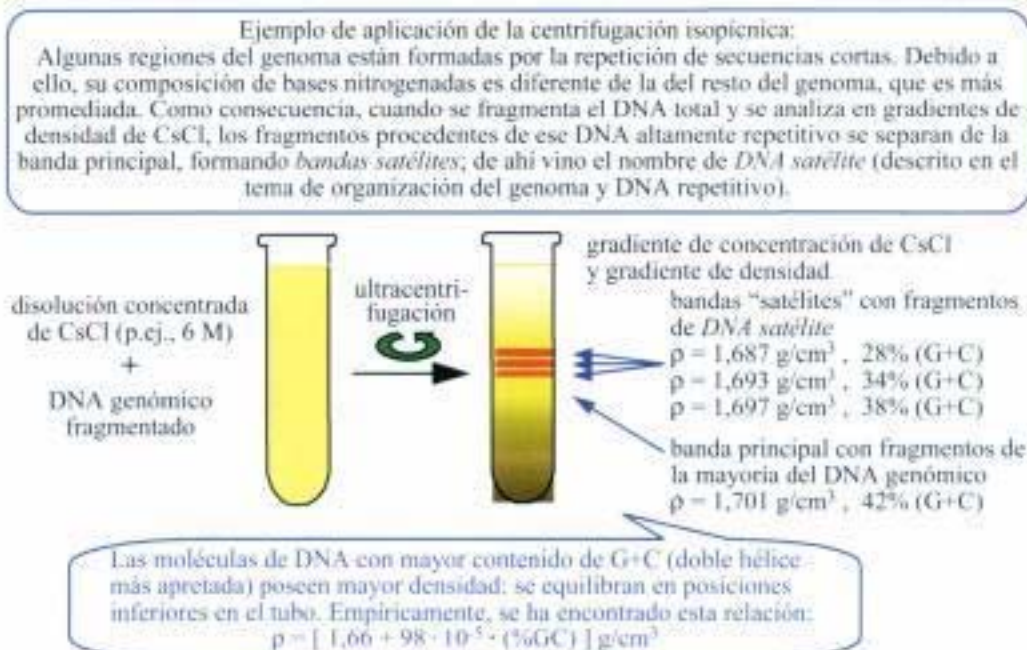
Este método se basa en el aumento de la **velocidad de sedimentación** por campos centrífugos lo suficientemente altos (superior a 100.000g) como para superar las fuerzas de difusión. Se emplea en especial para el DNA, tanto con fines preparativos (gran cantidad de material) como analíticos (pequeña cantidad de material).

11.12.1.1 Centrifugación zonal

De forma similar a lo estudiado para la separación de tipos celulares, se puede emplear esta modalidad de centrifugación sobre gradientes de densidad para separar moléculas de ácido nucleico, en función de su velocidad de sedimentación (definida por su coeficiente de sedimentación, en unidades Svedberg, S). Se consigue que las moléculas de ácido nucleico sedimenten a velocidad constante, con lo que se sitúan en bandas, cuya posición depende del tamaño y forma de la molécula. Estas se pueden recoger separadas en fracciones (por ejemplo, por punción del fondo del tubo) y cuantificarse midiendo su absorción a 260 nm.

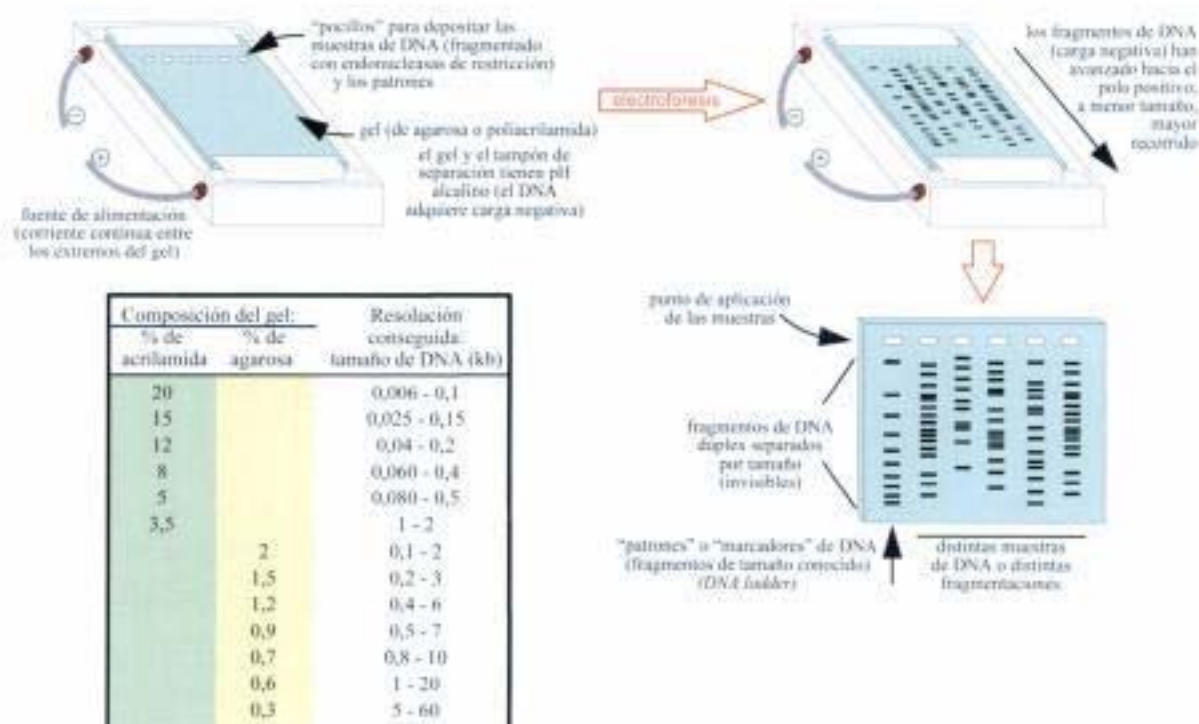
11.12.1.2 Centrifugación isopícnica

El método isopícnico o de equilibrio de sedimentación es quizá el más utilizado para la separación de distintos tipos de ácido nucleico, en especial DNA. Recuérdese que, en este caso, las macromoléculas se separan en función de su densidad, generalmente usando un gradiente que se va formando durante la ultracentrifugación a la vez que se separan los componentes de la muestra (gradiente *autoformado*). El medio habitual para preparar el gradiente es una disolución de cloruro de cesio que, gracias a su alta solubilidad y baja viscosidad, permite formar gradientes de un amplio margen de densidad. La muestra se mezcla con dicha disolución de CsCl y se procede a la ultracentrifugación durante largo tiempo (horas o días), hasta alcanzarse el equilibrio de sedimentación. Cada tipo de ácido nucleico se sitúa, formando una *banda*, en aquella posición del gradiente donde la densidad del medio iguala a la de la molécula.



11.12.2 Electroforesis

Como es bien sabido, esta técnica se basa en la capacidad de las macromoléculas cargadas para desplazarse en un campo eléctrico, con velocidad proporcional a su carga e inversamente proporcional a su tamaño. Las moléculas de DNA poseen, al pH alcalino habitual, una carga negativa uniforme por unidad de masa, lo que hace que su movilidad



hacia el ánodo (+) sólo venga determinada por el tamaño de la molécula (número de pares de bases, si es lineal). Debe destacarse que los cromosomas eucariotas son demasiado grandes para ser identificados por este tipo de análisis, por lo que comúnmente se aplica sobre DNA fragmentado con nucleasas.

11.12.2.1 Soporte para la electroforesis

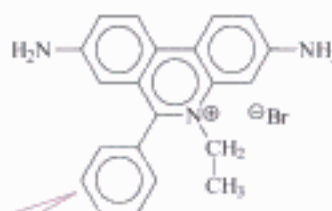
La separación de los fragmentos de DNA doble se realiza por electroforesis en geles de agarosa (generalmente en horizontal) o poliacrilamida (comúnmente en vertical). En ambos casos, el gel se prepara con distintas concentraciones en función del tamaño de los fragmentos de DNA que se quieren separar. Los fragmentos más pequeños (aproximadamente hasta 1.000 pb) se pueden separar adecuadamente en geles de poliacrilamida; para tamaños algo mayores (hasta 20 kpb) se usan geles de agarosa. Modificando la concentración de acrilamida o agarosa se consigue variar el intervalo de tamaños que es posible separar (resolución).

11.12.2.2 Detección del DNA tras su separación ("revelado")

Una vez separadas, las moléculas de DNA se visualizan normalmente por la fluorescencia de distintos compuestos que se unen a ellas. Aunque el más utilizado es el bromuro de etidio, va siendo sustituido por otros fluorocromos con mayor poder de detección o menor toxicidad.

Bromuro de etidio:

Agente intercalante, se introduce entre las bases apiladas del DNA, abre un poco la doble hélice, provoca un desenrollamiento local. Su fluorescencia aumenta mucho cuando se une al DNA (color naranja al iluminar con UV); método de tinción o revelado del DNA, especialmente usado en electroforesis y en centrifugación.

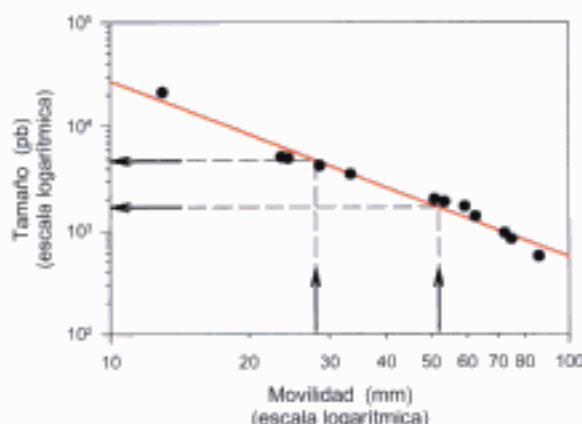
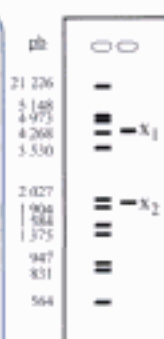


El bromuro de etidio debe manejarse con precaución porque posee carácter mutagénico.

11.12.2.3 Calibración de la electroforesis para el cálculo del tamaño de los fragmentos de DNA

Las moléculas lineales de DNA doble migran en la electroforesis a una velocidad inversamente proporcional a su tamaño (dentro del intervalo de resolución del gel empleado). Esto permite estimar la longitud de los fragmentos estudiados.

Se aplica en uno de los pocillos del gel una mezcla de fragmentos de DNA de tamaño conocido, denominados "patrones", "marcadores de peso molecular" o "escalera de DNA" (*DNA ladder*). Tras su separación en la electroforesis, se revelan y se representa para todas las bandas el logaritmo del tamaño (longitud en pb) frente al logaritmo de la movilidad (distancia de avance). Sobre la recta ajustada se interpolan las distancias de avance de las muestras problema, calculando así su tamaño en pb. En este ejemplo, los patrones empleados son fragmentos bien conocidos, aquellos obtenidos a partir del DNA del virus λ por la acción de las endonucleasas de restricción EcoRI y HindIII.



11.12.2.4 Electroforesis de campo pulsante

Las moléculas o fragmentos de DNA de longitud superior a 20-40 kpb no se pueden resolver (separar) empleando la electroforesis convencional en gel de agarosa. Esto se debe, en primer lugar, a la dificultad de manejo de los geles preparados con las bajas concentraciones de agarosa que requerirían (inferior al 0,4%). Además, las moléculas de DNA, que adoptan una conformación más o menos globular, poseen un tamaño excesivamente grande para atravesar los poros del gel y no se separan en función de su tamaño. Para solventar este problema se ha ideado una modificación de la técnica, conocida como **electroforesis de campo pulsante** (o *pulsado*, *pulsed-field gel electrophoresis*, PFGE).

Se emplean geles de agarosa al 1%, pero se altera periódicamente la orientación del campo eléctrico aplicado, activando alternativamente dos pares de electrodos. El frecuente cambio de dirección del campo eléctrico consigue que las moléculas se desplieguen y avancen a través de los poros en conformación extendida. A mayor tamaño, se reorientan con más dificultad, por lo que avanzan más despacio. Así se consigue separar fragmentos de DNA de hasta varias megabases, lo que supone un avance significativo pero aún no sirve para estudiar los cromosomas humanos enteros (de 50 a 250 Mb cada uno)



Las muestras han de contener el DNA sin fragmentar, por lo que no sirven las de DNA purificado, que generalmente ya se ha fragmentado por debajo de 100 kb como consecuencia de las fuerzas de cizalla durante los métodos habituales de aislamiento. Para obtener preparaciones con los cromosomas intactos se mezclan las células con agarosa caliente (37 °C) y se vierte en pequeños moldes de unos milímetros, de forma que al solidificarse la agarosa (4 °C) forma bloques ("insertos") que incluyen las células intactas. En éstos se realiza la lisis celular y digestión de las proteínas (p.ej., con EDTA, SDS y proteinasa K, que difunden a través de los poros del gel) sin alterar las grandes moléculas de DNA. Tras una diálisis para eliminar los restos de la digestión, se consigue un bloque de agarosa con el DNA intacto en su interior, que se deposita en un pocillo del gel de electroforesis.

11.12.3 Métodos cromatográficos

Aunque ampliamente utilizados en todos los campos, estos métodos sólo son apropiados para separar ácidos nucleicos en algunas aplicaciones. Por ejemplo, en columnas rellenas de *hidroxiapatito* se retiene por adsorción el DNA de doble hebra, y con menor afinidad el de hebra sencilla.

La *cromatografía de intercambio iónico*, utilizada clásicamente para separar proteínas, se ha convertido en un método alternativo para preparar DNA de alta pureza, de forma más rápida que la extracción orgánica o la centrifugación en gradiente de CsCl. Las columnas con resinas de intercambio aniónico (cargas positivas) retienen el DNA (cargado negativamente) dejando pasar otras moléculas (RNA, proteínas, etc.).

También se emplea la *cromatografía de afinidad* para fraccionar o separar RNAs de distintos tamaños, estructura y función. Por ejemplo, para purificar mRNA a partir del RNA celular total, se utilizan columnas de oligo-dT-celulosa o de poli-U-Sepharosa, que retienen el mRNA a través de la cola de poli(A) de este ácido nucleico, de forma similar a lo visto en el ejemplo de extracción directa con microesferas magnéticas.

Tema 12

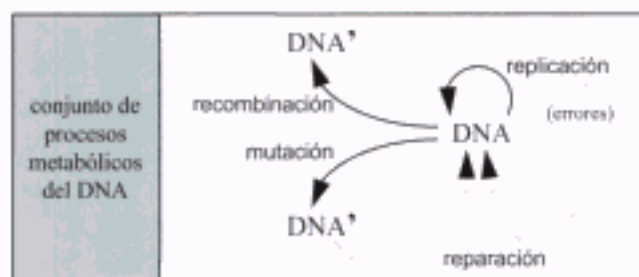
Replicación del DNA

12.1 Características generales de la replicación	141
12.1.1 Carácter semiconservador	142
12.1.1.1 Demostración experimental de la replicación semiconservadora	142
12.1.2 Síntesis simultánea, secuencial y bidireccional	144
12.1.3 Inicio monofocal o multifocal	145
12.2 Enzimología de la replicación	145
12.2.1 Requerimientos de la reacción de síntesis de DNA	145
12.2.2 Mecanismo de la reacción	146
12.1.3 Dirección de la síntesis	146
12.1.4 Tipos de DNA polimerasas	147
12.1.4.1 DNA polimerasas procarióticas: actividades polimerasa y exonucleasas	147
12.1.4.2 DNA polimerasas eucarióticas: actividades polimerasa y exonucleasa	148
12.1.5 Velocidad de replicación	149
12.3 Etapas en el proceso de replicación	150
12.3.1 Iniciación	150
12.3.1.1 Orígenes de replicación	150
12.3.1.2 Proteínas de iniciación	151
12.3.1.3 Conexión entre la iniciación y la elongación: proteínas necesarias	151
a) Helicasas	151
b) Proteínas de unión a hebra sencilla	151
c) Topoisomerasas	151
d) Primasa	152
12.1.2 Elongación	152
12.1.1.1 Asimetría de la replicación en ambas hebras	152
12.1.1.2 Mecanismo de la elongación	153
12.1.1.3 Maduración de los fragmentos de Okazaki	156
12.1.2 Terminación	157
12.1.2.1 Final de la elongación	157
12.1.2.2 Replicación de los telómeros	157
12.1.2.3 Telomerasa: componentes y características	158
12.1.2.4 Mecanismo de acción de la telomerasa	159
12.1.2.5 Funcionalidad de la actividad telomerasa en las células	160
12.2 Replicación mitocondrial	160

12.1 CARACTERÍSTICAS GENERALES DE LA REPLICACIÓN

La replicación es el proceso, conceptualmente sencillo pero molecularmente complejo, mediante el cual a partir de una molécula de DNA *progenitora* o *parental* se sintetiza una nueva, originándose así dos moléculas de DNA *hijas*, de secuencia idéntica a la del DNA original. Esto permite el paso de la información genética a la descendencia (tanto de una célula como del individuo).

Aunque la replicación constituye el aspecto esencial del **metabolismo del DNA**, también se deben considerar como parte del mismo otros procesos directa o indirectamente relacionados: la **recombinación** o reordenamiento de la información genética, la **mutación** o alteración de la secuencia y la **reparación** de las alteraciones o daños en el DNA.



La replicación se produce de forma coordinada con la división celular (tema 9), concretamente en la fase S, durante las interfases previas tanto a la mitosis como a la meiosis I, de forma que, en el caso de la mitosis, las dos células hijas reciben la misma dotación genética que tenía la célula madre. En cuanto a la meiosis, puesto que da lugar a cuatro células hijas, la replicación previa permite que éstas reciban la mitad de dicha dotación genética, y no una cuarta parte (pág. 104). La relación entre el estado de replicación y la progresión por el ciclo celular se puede resumir en los siguientes principios generales:

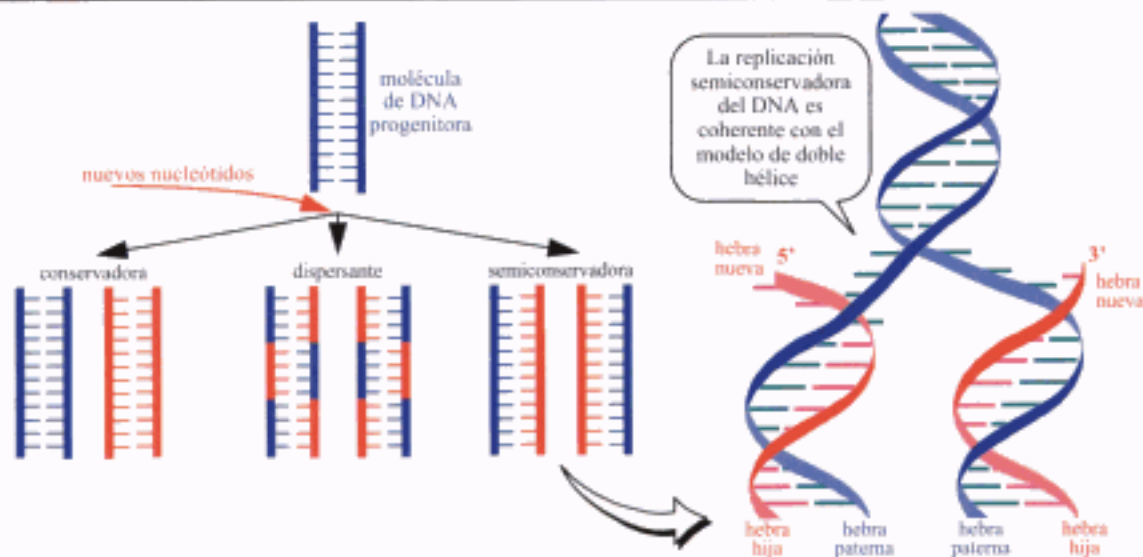
- El inicio de la replicación del DNA obliga a la célula a emprender una división.
- Una vez iniciada la replicación, no puede tener lugar la consiguiente división hasta que se haya completado dicha replicación. En realidad, el final de la replicación constituye posiblemente una señal para la división.
- El proceso de replicación se lleva a cabo sin necesidad de una separación completa de la cadena progenitora: la doble hélice se desenrolla gradualmente y sus dos hebras se van separando a la par que se produce su replicación.

Antes de estudiar los detalles moleculares de la síntesis de un nuevo DNA, procede analizar las características principales del proceso, atendiendo a un criterio eminentemente didáctico. Éstas son: carácter **semiconservador**, realización **simultánea** en ambas hebras, de forma **secuencial** y con carácter **bidireccional**, y **origen** monofocal (procariotas) o multifocal (eucariotas). Otra característica destacada, el carácter **semidiscontinuo**, se explicará una vez que se conozca el mecanismo de la reacción.

12.1.1 Carácter semiconservador

Aunque teóricamente son posibles **tres mecanismos** para la replicación, dos de ellos son hipotéticos y pronto se demostró que sólo el tercero tiene lugar en las células:

<i>Replicación conservadora</i>	<i>Replicación dispersante</i>	<i>Replicación semiconservadora</i>
Un complejo enzimático recorrería el DNA, reconociendo la secuencia en ambas hebras para copiarlas simultáneamente, de forma que el DNA progenitor de doble hebra se conservaría intacto mientras que el DNA hijo contendría dos hebras totalmente nuevas.	El DNA progenitor se rompería varias veces antes de que se sintetizaran los nuevos fragmentos de DNA. Éstos se unirían al DNA progenitor para dar lugar a los dos DNAs hijos, ambos con fragmentos tanto nuevos como de las dos hebras del DNA progenitor.	Es el mecanismo real. Las dos hebras del DNA progenitor sirven de molde para la síntesis de sus respectivas hebras complementarias. Cada DNA hijo está formado por una hebra progenitora y otra hija (recién sintetizada). El mecanismo se repite en sucesivas generaciones de células. El nombre alude, precisamente, a que siempre se conserva la mitad de la molécula original.



12.1.1.1 Demostración experimental de la replicación semiconservadora

El mecanismo real de la replicación fue demostrado por primera vez por Mathew Meselson y Franklin W. Stahl en 1958, mediante experimentos muy elegantes, sencillos y representativos de muchos otros realizados en diversas áreas de la Biología Molecular. Aunque los aspectos moleculares de la replicación fueron establecidos posteriormente con el estudio de la DNA polimerasa, el experimento de Meselson y Stahl tiene interés, aparte de por su propio diseño, por haber sido la primera vez que se aplicaron tanto la técnica de sedimentación en gradiente de densidad como la introducción de isótopos en las moléculas de DNA (^{15}N , infrecuente, y ^{14}N , isótopo habitual).

Para comprender la novedad y trascendencia de estos experimentos en su momento y contexto, se van a considerar los antecedentes y el planteamiento antes de describir el resultado. Se aprovecha así este ejemplo para ilustrar el método científico, en especial cómo un diseño adecuado del experimento, que parece simple, permite obtener una información notable.

a) Antecedentes

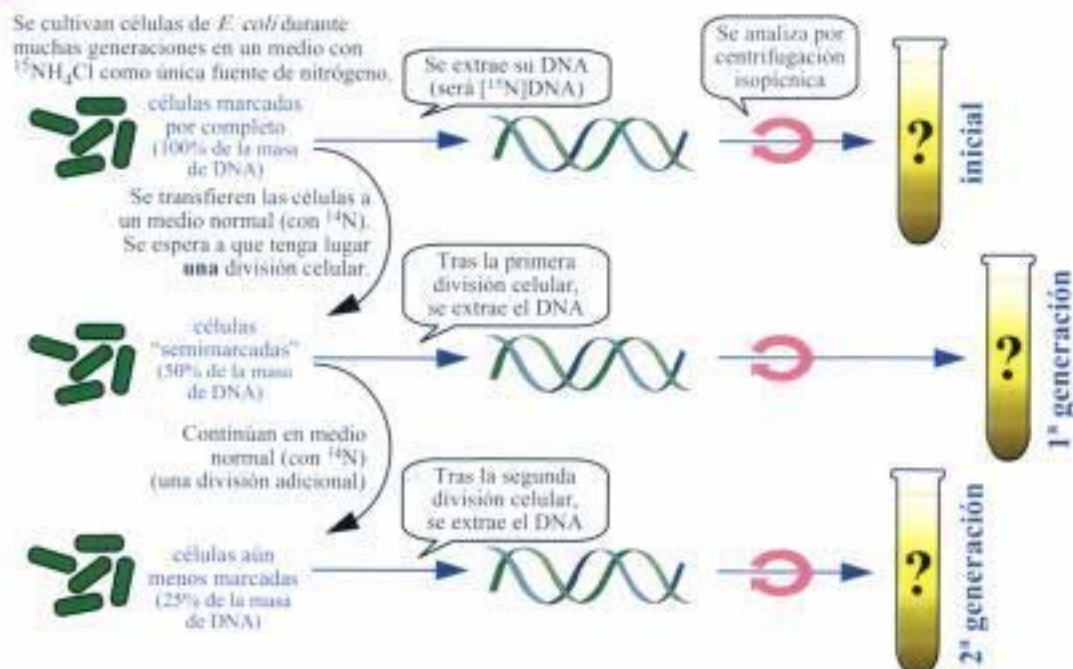
El primer lugar, se examina la posibilidad de aplicar la técnica de equilibrio de sedimentación en gradiente de densidad, o centrifugación isopícnica (pág. 124), al análisis de muestras de DNA cuya densidad es diferente por estar marcado o no con el isótopo ^{15}N .



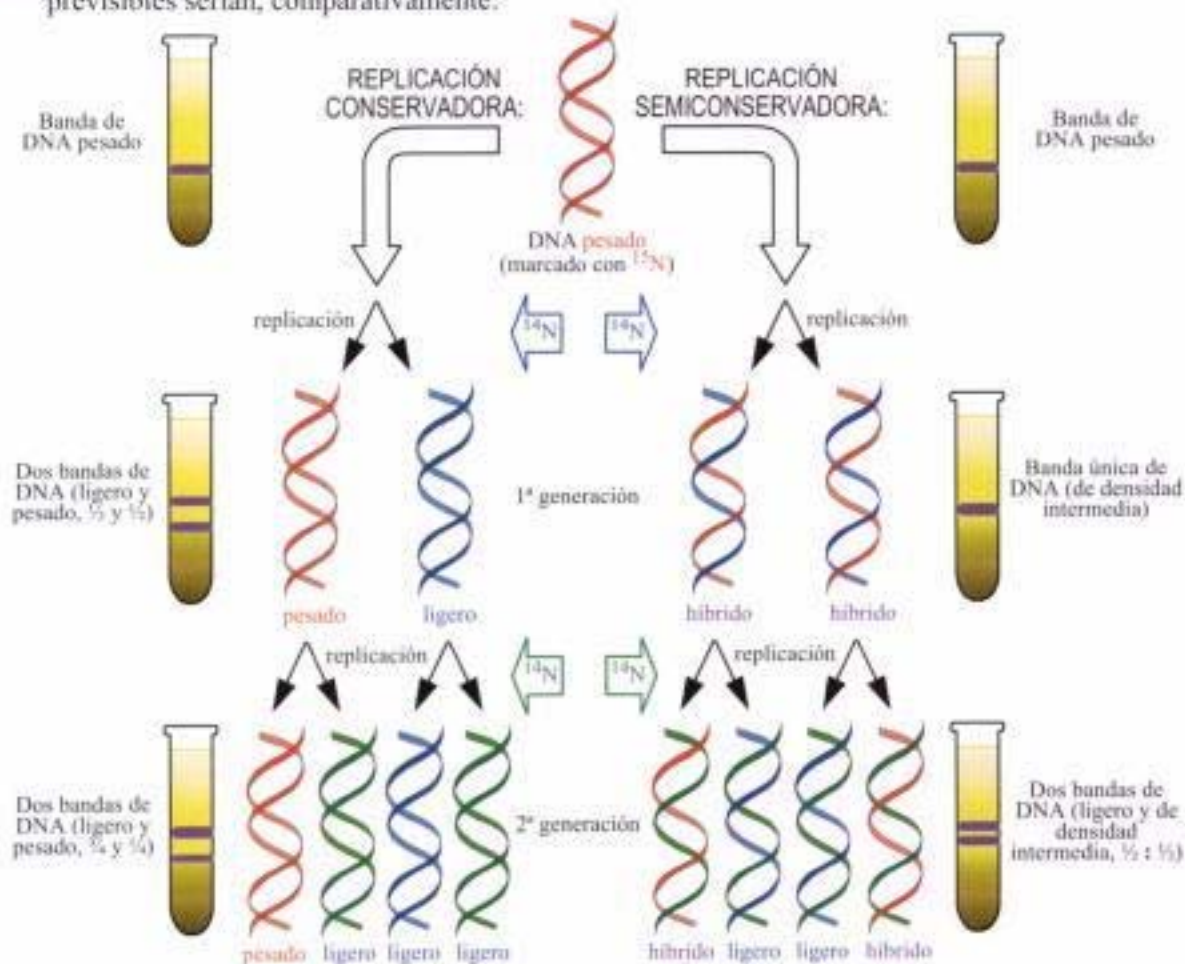
b) Planteamiento y diseño del experimento

Aplicando esta técnica de marcaje del DNA con ^{15}N y su análisis por centrifugación isopícnica, se diseñó el experimento de forma que los resultados obtenidos permitieran diferenciar cuál de las hipótesis era la correcta:

① Diseño del experimento:



- 2) Bajo las hipótesis de replicación conservadora y semiconservadora, los resultados previsibles serían, comparativamente:



Aunque no se presentan los resultados, la REPLICACIÓN DISPERSANTE habría producido moléculas de DNA con mezcla aleatoria de nucleótidos progenitores y nuevos, es decir, en la 1ª generación una sola banda en el gradiente (densidad intermedia entre pesado y ligero) y en la 2ª también una sola banda (densidad intermedia entre la anterior y la del ligero).

c) Resultados obtenidos y conclusión

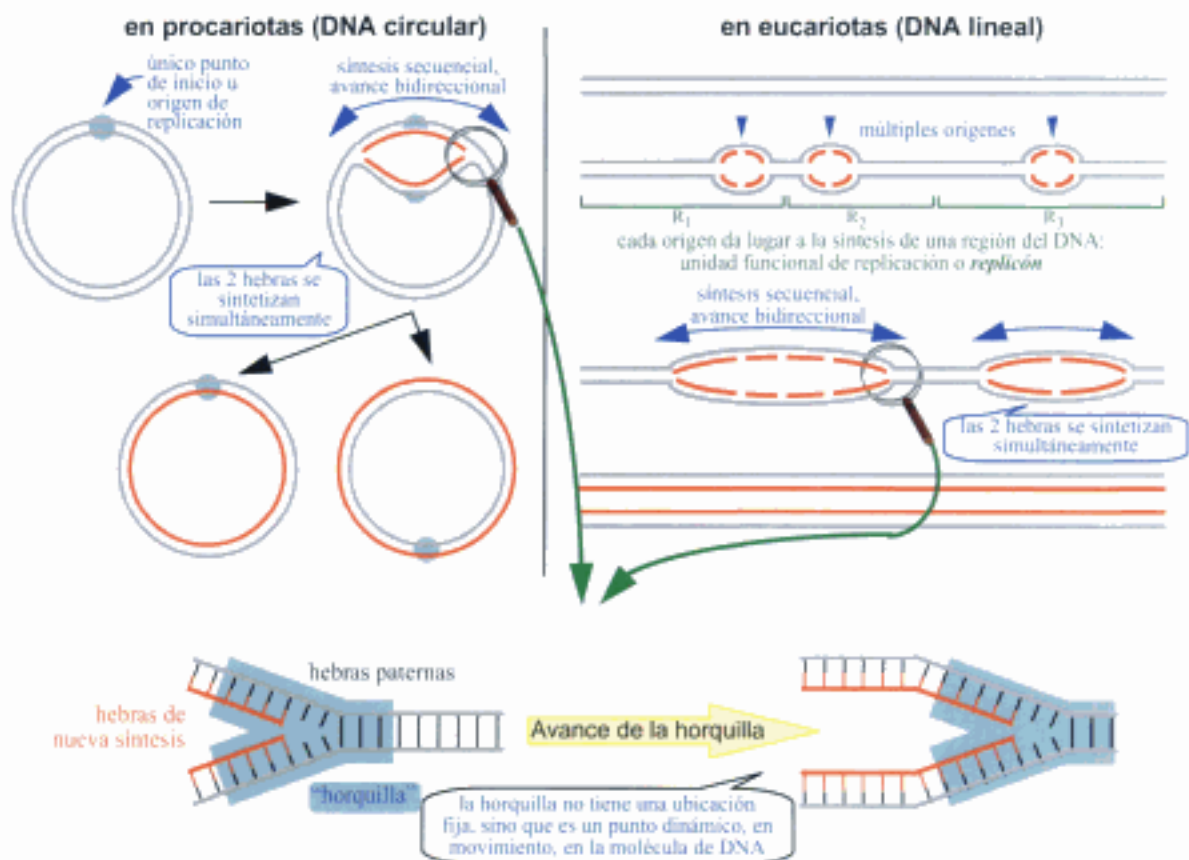
El resultado observado al realizar el experimento fue de una banda de DNA de densidad intermedia en la primera generación de células y dos bandas en la segunda generación, de densidades baja e intermedia. Al coincidir los datos experimentales con los resultados previsibles de la parte derecha de la figura, se concluyó que **la hipótesis correcta es la semiconservadora**, es decir, en la replicación del DNA cada una de las moléculas hijas conserva una hebra del DNA progenitor y tiene otra hebra de nueva síntesis.

12.1.2 Síntesis simultánea, secuencial y bidireccional

De cara a la descripción del mecanismo de la replicación es importante destacar algunas características, observadas por primera vez en *E. coli* pero igualmente válidas para eucariotas (mediante la adición al medio de cultivo de timidina tritiada para marcar el DNA de nueva síntesis y luego detectarlo por métodos autorradiográficos).

- La síntesis no comienza al azar, sino en puntos concretos de la molécula, denominados *orígenes de replicación*. En cada uno de ellos se produce la apertura local de la doble hélice y comienza la *síntesis simultánea* de las dos nuevas hebras (no debe confundirse *simultáneo* con el carácter *continuo* de la síntesis de una sola de las hebras, que se estudiará más adelante).
- Como consecuencia, las dos hebras nuevas se van alargando progresivamente, por adición *secuencial* de nucleótidos.
- La separación de las hebras progenitoras, que comienza en cada origen, progresa en ambas direcciones, por lo que se dice que la replicación es *bidireccional*. A los puntos de transición entre doble hebra y hebras sencillas se los llama *horquillas de replicación*, y van alejándose entre sí.
- Finalmente, no debe pasarse por alto el hecho de que, dado que las dos hebras del DNA son antiparalelas, su separación con el avance de la horquilla progresa en sentido opuesto, para una hebra de 3' hacia 5' y para la otra en sentido 5' $\rightarrow$ 3' (a la inversa en la otra horquilla).

FORMACIÓN Y AVANCE DE LAS HORQUILLAS DE REPLICACIÓN



12.1.3 Inicio monofocal o multifocal

En el cromosoma circular único de procariotas la replicación comienza siempre en un punto determinado, denominado origen (y marcado como ori^C); se dice que la replicación es **monofocal**. En consecuencia, progresa formando dos horquillas de replicación. Por el contrario, en eucariotas la replicación es **multifocal**, pues en cada uno de los cromosomas existen múltiples orígenes de replicación (centenares o miles), que dan lugar a un número doble de horquillas. Gracias a ello se hace posible completar en un tiempo razonable la replicación de los cromosomas, mucho más largos que en procariotas.

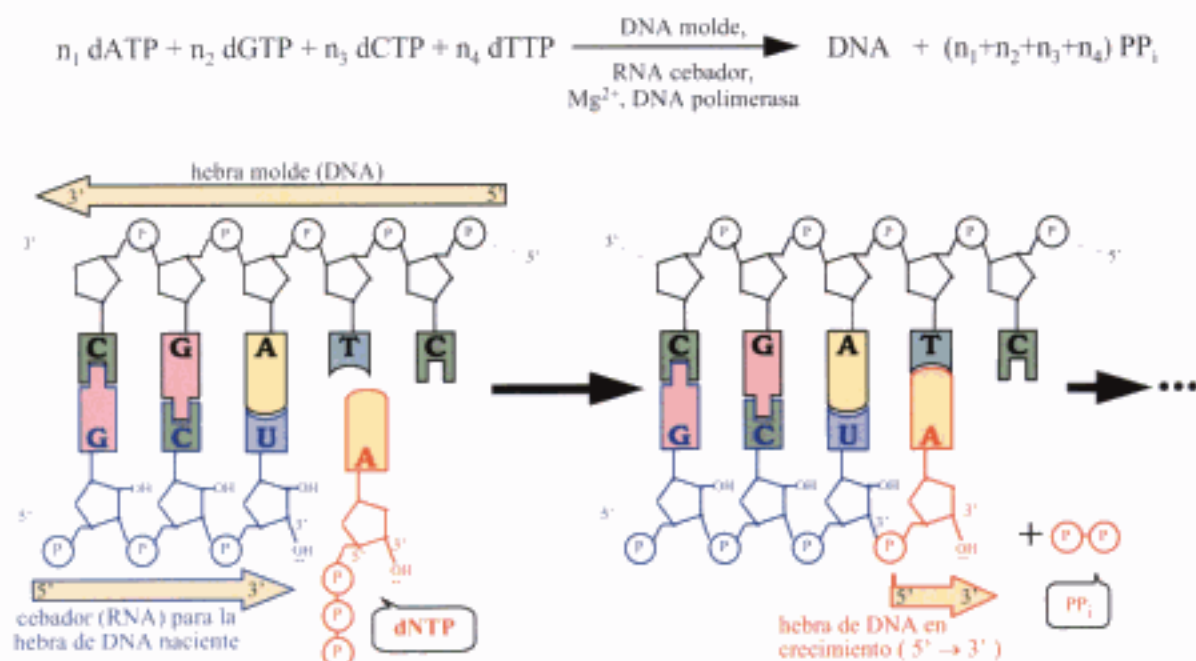
12.2 ENZIMOLOGÍA DE LA REPLICACIÓN

12.2.1 Requerimientos de la reacción de síntesis de DNA

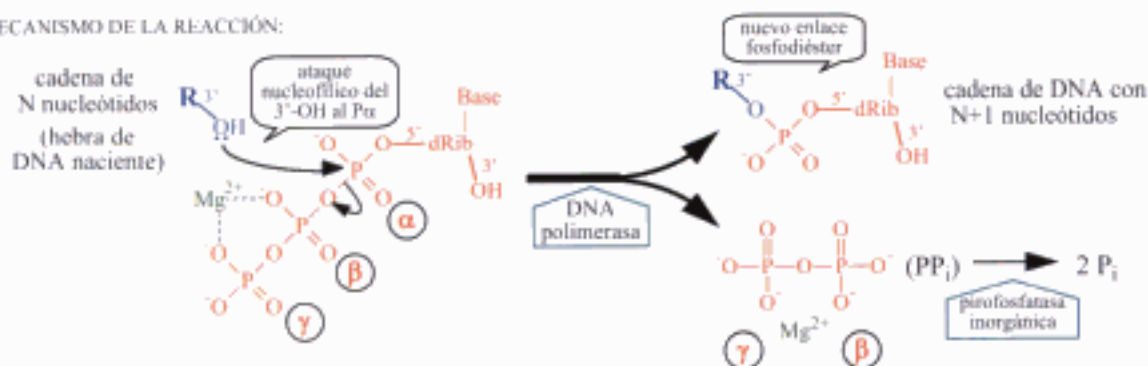
- **Cebador:** La síntesis de una nueva hebra de DNA no puede comenzarse a partir de dos nucleótidos, sino que requiere un fragmento de hebra iniciador, denominado cebador, que aporte un grupo 3'-OH libre. Se trata de un oligonucleótido, por lo común RNA, que debe estar apareado a la hebra progenitora de DNA, de forma complementaria y antiparalela. Por tanto, se puede decir que la replicación no es autoiniciadora (a diferencia de la transcripción, pág. 248), sino que sólo **elonga** o alarga moléculas preexistentes.
- **Sustratos:** Se utiliza como sustratos el conjunto de **los cuatro dNTPs**: dATP, dGTP, dCTP y dTTP. De cada uno de ellos queda incorporado en el DNA nuevo la parte dNMP de la molécula. A estos efectos, los dNMPs y dNDPs son inactivos.
- **Cofactores:** Para una actividad óptima se requiere un ion metálico divalente como cofactor, asociado a los dNTPs. Aunque *in vitro* este papel pueden desempeñarlo tanto Mn^{2+} como Mg^{2+} , es este último el que actúa *in vivo*.
- **Molde o plantilla:** Ésta es, quizá, la principal característica de la reacción. El orden correcto de incorporación de los dNMPs viene determinado por su complementariedad de bases con la secuencia de cada hebra de DNA, que actúa así como *molde o plantilla*.

12.2.2 Mecanismo de la reacción

Se inicia la reacción con el apareamiento de un dNTP complementario al nucleótido de la hebra molde (DNA) en la posición vecinal al extremo 3' de la hebra en crecimiento (inicialmente, el RNA cebador; luego, la hebra que se está sintetizando). La reacción propiamente dicha consiste en la unión del dNTP seleccionado, bajo la forma de dNMP, con liberación de PP_i . Ello implica la formación de un puente fosfodiéster entre el fosfato α del dNTP que se incorpora y el 3'-OH libre de la cadena en crecimiento (inicialmente, del cebador). La ruptura del enlace fosfoanhidrido (α - β) del dNTP proporciona la energía necesaria para impulsar la reacción, ayudada por la subsiguiente hidrólisis del PP_i a $2 P_i$ catalizada por la pirofosfatasa. Tras muchas reacciones similares sucesivas, el producto es la nueva hebra de DNA:



MECANISMO DE LA REACCIÓN:



12.2.3 Dirección de la síntesis

Puesto que la **adición de nucleótidos** se realiza siempre a partir de 5'-trifosfatos y sobre un extremo 3'-OH libre, la hebra crece por su extremo 3', y se dice que la síntesis transcurre en **dirección 5'→3'**. Esto es válido para todas las DNA polimerasas, tanto de procariotas como de eucariotas (y también para las RNA polimerasas).

12.2.4 Tipos de DNA polimerasas

Aunque la reacción de síntesis de DNA es universal, es decir, tiene lugar de idéntica forma en todos los organismos y ubicaciones subcelulares, su catálisis es efectuada en cada caso por distintas enzimas, bajo la denominación común de **DNA polimerasas** (abreviadas como **DNApol**). Estrictamente su nombre es más completo, **polimerasas de DNA dirigidas por DNA** (o **dependientes de DNA**), para indicar tanto la molécula que sintetizan como la que emplean como molde. Más adelante se estudiarán otras polimerasas de DNA o de RNA que utilizan distintos moldes; para ofrecer una visión unificada se resumen a continuación:

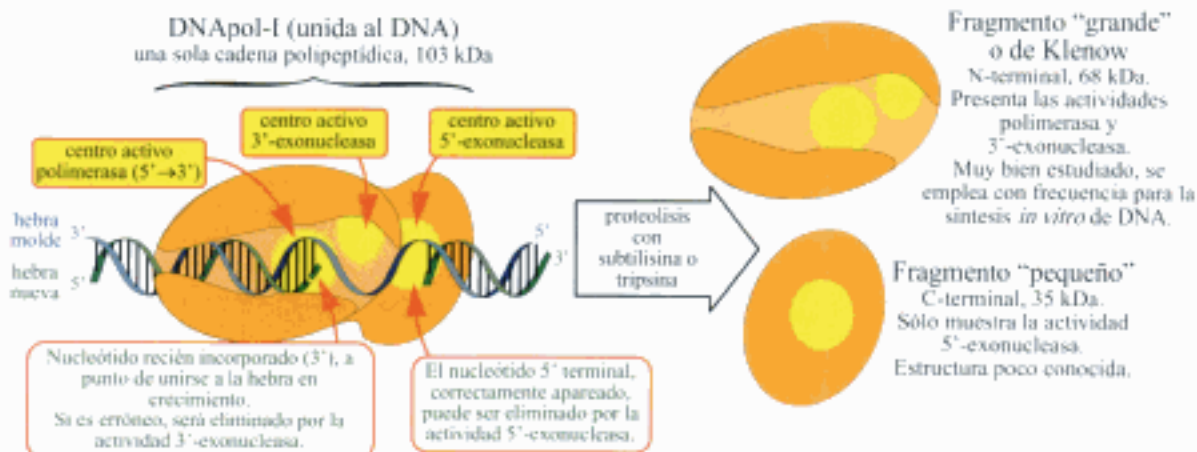
molécula sintetizada	molécula molde	nombre completo	ejemplos (nombre común)
DNA	DNA	polimerasa de DNA dirigida por DNA	• DNA polimerasas
RNA	DNA	polimerasa de RNA dirigida por DNA	• RNA polimerasas o transcriptasas
			• primasas
DNA	RNA	polimerasa de DNA dirigida por RNA	• transcriptasas inversas
			• telomerasas
RNA	RNA	polimerasa de RNA dirigida por RNA	• RNA replicasas

12.2.4.1 DNA polimerasas procarióticas: actividades polimerasa y exonucleasas

La primera DNA polimerasa descubierta (Arthur Kornberg, 1958) y la mejor estudiada hasta la fecha es la **DNApol-I** de *Escherichia coli*. La comprobación de que su actividad no es suficiente para justificar la velocidad de replicación del cromosoma bacteriano *in vivo* condujo al descubrimiento de otras enzimas: las **DNApol-II** y **DNApol-III**. Las tres coexisten en *E. coli*, con funciones especializadas: pol-III realiza la mayor parte de la tarea de replicación, mientras que pol-I y pol-II se encargan de funciones auxiliares (sus características particulares se muestran en la tabla de pág. 148, junto con las de las polimerasas eucarióticas).

Las tres DNA polimerasas presentan, además de su actividad enzimática como polimerasa, una o dos **actividades exonucleasa**. La actividad **3'-exonucleasa** (también llamada **exonucleasa 3'→5'**) permite hidrolizar el nucleótido situado en el extremo 3' de la hebra de DNA, sólo un nucleótido cada vez. Esta eliminación tiene lugar cuando el nucleótido incorporado por la actividad polimerasa a la cadena en crecimiento es erróneo, es decir, no aparea correctamente con el presente en la hebra molde de DNA. La polimerasa es así capaz de reconocer el error, detener la adición de más nucleótidos y liberar ese último dNMP mal incorporado. Los centros activos polimerasa y exonucleasa son diferentes. En resumen, las DNA polimerasas corrigen sus propios errores, en sentido (3'→5') contrario al de la polimerización (5'→3'); por ello, a esta actividad 3'-exonucleasa se la llama también **actividad correctora de pruebas** (o de galeras, por analogía al proceso en imprenta), o correctora de errores propios. De este modo mejora la fidelidad de la replicación, en un factor de entre 10^2 y 10^3 sobre la que aporta la incorporación de nucleótidos complementarios por la polimerasa (que tiene una tasa de error de un nucleótido por cada 10^4 o 10^5).

CARACTERÍSTICAS DE LA DNA POLIMERASA I DE *E. coli*



La DNApol-I de procariotas es única entre las polimerasas por poseer además una segunda actividad exonucleasa, en este caso **5'-exonucleasa** (o **exonucleasa 5'→3'**), que le permite hidrolizar secuencialmente DNA o RNA a partir del extremo 5' de la hebra. De nuevo, el centro activo es independiente de los anteriores; en este caso está situado en el fragmento pequeño. Esta actividad juega un papel fundamental en la replicación del DNA procariótico, al eliminar el RNA cebador y sustituirlo por DNA, durante la síntesis discontinua de la hebra retardada (pág. 156). También

complementa la actividad 3'-exonucleasa corrigiendo errores de distinto tipo, y se la llama en ocasiones **actividad de reparación**; por ejemplo, interviene en la escisión de los dímeros de pirimidina que se forman por exposición del DNA a luz UV. Experimentalmente, se aprovecha la actividad 5'-exonucleasa, por ejemplo, para la preparación de sondas marcadas (pág. 182).

Algunas características de las DNA polimerasas, desde el punto de vista bioquímico y funcional

Polimerasa:	PROCARIOTAS			EUCARIOTAS				
	I	II	III	α	β	γ	δ	ϵ
Ubicación subcelular				núcleo	núcleo	mitocondria	núcleo	núcleo
Actividades enzimáticas:								
Primasa (inicio)	no	no	no	sí	no	no	no	no
Polimerasa 5'→3' (elongación)	sí	sí	sí	sí	sí	sí	sí	sí
3'-Exonucleasa (o 3'→5') (corrección de pruebas)	sí	sí	sí	no	no	sí	sí	sí
5'-Exonucleasa (o 5'→3') (eliminación de cebadores)	sí	no	no	no	no	no	no	no
Función en la célula:								
Replicación (síntesis continuada de hebras nuevas, conductora y retardada)	no	no	sí	sólo inicialmente	no	sí	sí	sí
Empalme de fragmentos de Okazaki	sí	no	no	no	no		sí (con nucleasas)	sí
Velocidad de replicación:								
Velocidad de polimerización (nucleótidos/segundo)	16-20	2-7	250-1000					
Procesividad (nucleótidos añadidos antes de su disociación)	baja 3-200	media ≥10.000	muy alta ≥5 · 10 ⁵	baja 100-200	baja	elevada	elevada con PCNA ≥5.000	elevada con/sin PCNA

12.2.4.2 DNA polimerasas eucarióticas: actividades polimerasa y exonucleasa

En **eucariotas**, las DNA polimerasas son diferentes a las de procariotas, aunque presentan cierta similitud de secuencia y estructura. Se han descrito más de **cinco enzimas**. La **DNApol α** se encarga, gracias a su actividad primasa, de iniciar la síntesis mediante la formación de un RNA cebador y, por su actividad polimerasa, de la elongación inicial de ese cebador. Esta tarea se cede luego a las **DNApol δ y ϵ** , responsables de la mayor parte de la elongación de ambas hebras; de hecho, ambas tienen actividades similares, pero aún no está claro si desempeñan indistintamente la misma función o están especializadas. La **DNApol β** no interviene en la replicación, sino que está implicada en la reparación de errores o daños en el DNA (procesos en los que también participan otras polimerasas, incluyendo δ y ϵ). Mientras estas cuatro enzimas son responsables de la replicación, reparación y recombinación del DNA nuclear, el DNA mitocondrial es sintetizado por una polimerasa diferente, la **DNApol γ** . Por último, se han descubierto recientemente otras polimerasas nucleares, aún poco conocidas (ζ [zeta], η [eta], ι [iota]), que están implicadas sobre todo en reparación y recombinación.

En cuanto a la actividad exonucleasa de las DNA polimerasas eucarióticas, sólo las 3 que se encargan de la elongación (δ , γ y ϵ ; ver tabla) poseen actividad **3'-exonucleasa**, aquella responsable de la capacidad de corrección de errores cometidos al incorporar los nucleótidos a la hebra que se está sintetizando. Ninguna de las polimerasas eucarióticas parece tener actividad 5'-exonucleasa; la función de eliminación de cebadores la desempeñan dos nucleasas independientes, aunque las polimerasas δ y ϵ sí se encargan de rellenar el hueco con DNA; los detalles se estudian más adelante (pág. 156).

La función específica de cada polimerasa sólo se puede comprender en el contexto de los mecanismos de iniciación y elongación de las cadenas de DNA, que se estudian con detalle dentro del mecanismo de la elongación (pág. 154).

Bajo el alcance de este libro no parece adecuado describir las propiedades estructurales de las distintas polimerasas. En la tabla siguiente se resumen los datos relativos a tamaño y composición en subunidades, pero debe tenerse en cuenta que todavía es muy elevada la variabilidad de cifras, en especial para las enzimas eucarióticas, e incluso la interpretación de los mismos en los distintos organismos estudiados, como para su generalización a todos los eucariotas.

Características estructurales de las DNA polimerasas

Polimerasa:	PROCARIOTAS			EUCARIOTAS ⁽¹⁾				
	I	II	III	α	β	γ	δ	ϵ
M _r total (kDa) ⁽²⁾	103	90	~900	356	39	160	173	313
Nº subunidades	1	≥4	22	4	1	2	2 ó 3	2
M _r subunidades			⁽³⁾	⁽⁴⁾		125	125	258
						35	48	55

⁽¹⁾ Los datos corresponden a humanos.

⁽²⁾ Masa molecular relativa de la holoenzima, o enzima funcional con todas sus subunidades.

⁽³⁾ La DNApol-III de *E. coli* está formada por 22 subunidades de 10 tipos, con una composición global $(\alpha, \epsilon, \theta)_2 \tau_2 (\gamma, \delta, \delta', \psi, \chi)_2 \beta_3$. De ellas, la agrupación $(\alpha, \epsilon, \theta)$ se conoce como "núcleo" de la polimerasa. La subunidad α (130 kDa) tiene el centro activo polimerasa, la ϵ el 3'-exonucleasa o correctora de pruebas.

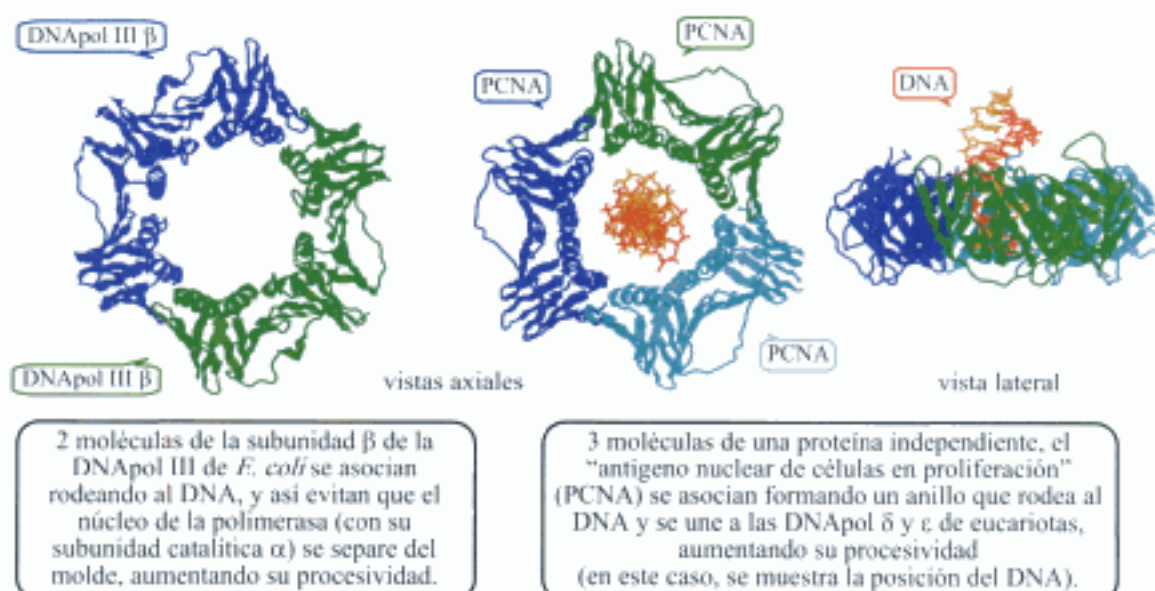
⁽⁴⁾ La DNApol α humana está formada por 4 subunidades diferentes: 180 kDa, con el centro activo polimerasa; 48 kDa, con el centro activo primasa; 58 kDa, cofactor de la primasa; 70 kDa, reguladora de la polimerasa.

12.2.5 Velocidad de replicación

Dado el gran tamaño del genoma de una célula, para que su replicación completa se lleve a cabo en un tiempo razonable el proceso de síntesis debe ser muy rápido. En procariotas, la velocidad llega a 1.000 nucleótidos por minuto. En eucariotas, el mayor tamaño del genoma hace pensar en la necesidad de una gran velocidad de replicación que, sin embargo, no precisa ser tan elevada gracias a la existencia de varios orígenes de replicación en cada cromosoma, que reduce la longitud de DNA que debe sintetizar cada molécula de polimerasa.

La consecución de una alta velocidad de síntesis sólo es posible gracias a un mecanismo peculiar, según el cual la molécula de DNA polimerasa se mantiene unida a la cadena de DNA molde, a la vez que asociada por su centro activo al extremo 3'-OH de la hebra creciente. De esta forma puede incorporar numerosos nucleótidos sucesivamente, con una reducción del tiempo necesario para la adición de cada uno: por ejemplo, la polimerasa puede tardar un minuto en soltarse y volverse a unir al DNA (por mera difusión de las moléculas), mientras que para añadir un nucleótido sin soltarse le bastan unos milisegundos. Esta capacidad de las polimerasas, variable entre ellas, se recoge bajo el término *procesividad*, que se define como el número de nucleótidos que son incorporados a la hebra antes de que la enzima se separe del molde (coloquialmente se diría añadidos "de un tirón"). Un valor elevado de procesividad en una polimerasa es mucho más importante para la replicación que una velocidad alta de catálisis.

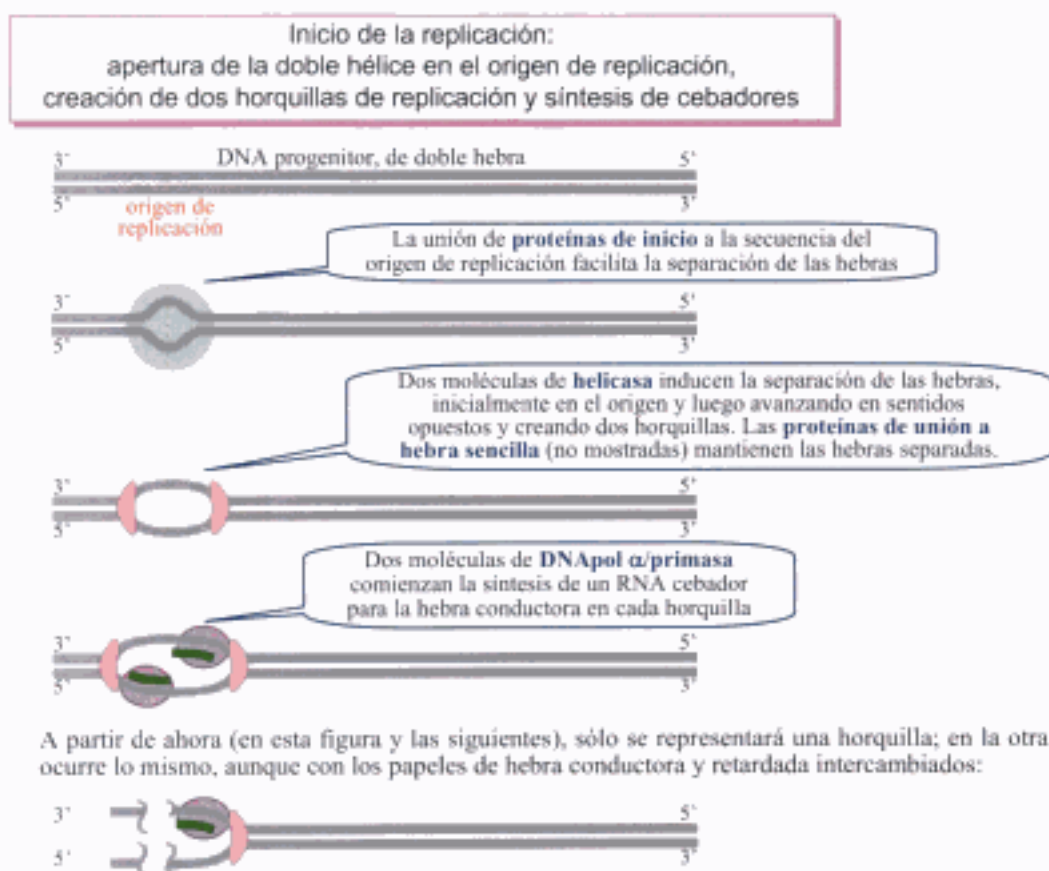
Las polimerasas de mayor procesividad son las que intervienen en la elongación de las hebras de DNA (tabla en pág. 148), mientras que las de baja procesividad se encargan de tareas que podrían llamarse auxiliares, tales como la síntesis y eliminación de cebadores, su sustitución por DNA o la reparación. Generalmente, la alta procesividad se consigue gracias a una molécula que "abrazo" o rodea la cadena de DNA, evitando así que la enzima se separe; este mecanismo se denomina de "pinza deslizante", pues permite el deslizamiento continuado de la polimerasa a lo largo de la cadena de DNA, sin separarse de ella.



12.3 ETAPAS EN EL PROCESO DE REPLICACIÓN

12.3.1 Iniciación

Mientras que en procariotas es probable que la replicación tenga lugar en conexión con la membrana celular, pues el único cromosoma está anclado a ella en la zona nucleóide, en eucariotas el proceso ocurre, lógicamente, en la matriz nuclear. El DNA se replica una vez por ciclo celular, durante la fase S.



12.3.1.1 Orígenes de replicación

La replicación siempre comienza en puntos de la molécula de DNA con secuencias características. En el caso de **procariotas**, el cromosoma tiene un origen de replicación único, llamado *ori^C*; por ejemplo, en *E. coli* está constituido por una secuencia de 245 pb. En **eucariotas**, como ya se ha comentado, la enorme longitud de los cromosomas requiere la existencia de numerosos orígenes de replicación; aunque peor conocidos, éstos poseen también secuencias características.

El proceso de replicación se realiza en la unidad funcional del genoma denominada **replicón**, definida como aquella región de DNA que puede ser replicada a partir de un mismo origen (pág. 145). Cada replicón contiene, pues, un origen de replicación y se replica mediante dos horquillas. Mientras en procariotas el cromosoma completo constituye un solo replicón, en eucariotas son numerosos y sufren la replicación de forma casi simultánea, lo que supone la existencia de un complejo mecanismo de control que coordine el inicio, elongación y terminación en todos los replicones de cada cromosoma. Estos aspectos están pendientes de resolución, al igual que muchos otros relacionados, por ejemplo, con la replicación de los centrómeros y otras regiones heterocromáticas altamente condensadas.

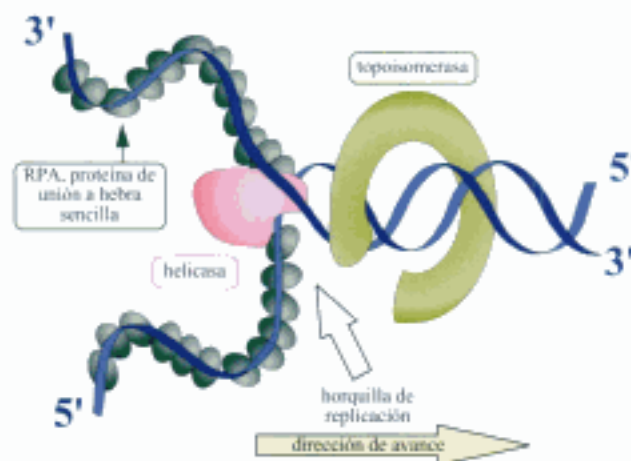
Centrándonos en el inicio y progreso de la replicación en cada replicón, surgen igualmente grandes cuestiones relativas a cómo son las interacciones específicas entre DNA y proteínas que desencadenan la iniciación, a la consiguiente actuación de estas proteínas y al control del proceso. Por delante de la horquilla de replicación, apenas se conoce cómo tienen lugar la necesaria descondensación de los cromosomas, la disociación de los nucleosomas y la separación de las proteínas histonas y no histonas (temas 8 y 9). Por detrás de la horquilla, debe también aclararse cómo se reutilizan las histonas y no histonas para la reasociación de los nucleosomas, cómo tiene lugar la lenta organización completa de la cromatina (que no parece establecerse hasta que la horquilla de replicación se encuentra muchas kb más adelante), etc.

12.3.1.2 Proteínas de iniciación

Tanto en procariotas como en eucariotas, el origen de replicación es reconocido por varias moléculas proteicas, que provocan el pliegue de la cadena de DNA y crean una tensión superhelicoidal negativa que facilita la separación de las hebras, generalmente ayudada por la abundancia de pares AT en esa región. Por ejemplo, en la secuencia *ori^C* esta acción se debe a la unión específica de un oligómero de la proteína denominada DnaA, que atrae a otras proteínas, principalmente DnaB (una helicasa) y DnaC. También existen proteínas de iniciación que se unen específicamente a las secuencias de inicio eucarióticas, formando un "complejo de reconocimiento del origen".

12.3.1.3 Conexión entre la iniciación y la elongación: proteínas necesarias

Una vez formado el complejo de iniciación, la replicación precisa el desenrollamiento de la doble hélice, a fin de hacer accesibles las bases como molde para formar las nuevas hebras. El desenrollamiento inicial en el tramo correspondiente al origen de replicación, por efecto de las proteínas de iniciación, debe continuar avanzando por delante de la polimerasa, encabezando la horquilla de replicación. Además, a medida que se van sintetizando las hebras nuevas se debe ir recuperando el enrollamiento en las nuevas dobles hélices. En estas operaciones de desenrollamiento o relajamiento y enrollamiento o compactación intervienen diversas proteínas especializadas. El complejo multiproteico formado por éstas y por las DNA polimerasas, que viaja asociado a cada horquilla y realiza la replicación, se conoce como **replisoma**.



a) Helicasas

Estas proteínas multiméricas (generalmente hexámeros) se unen al DNA en el origen y actúan separando sus dos hebras, en un proceso que requiere la hidrólisis simultánea de ATP. Se crean así dos horquillas de replicación, que comienzan a avanzar en sentidos opuestos, cada una liderada por una helicasa, que avanza por la cadena seguida inmediatamente por el resto de componentes del replisoma o complejo de elongación asociado a la horquilla.

b) Proteínas de unión a hebra sencilla

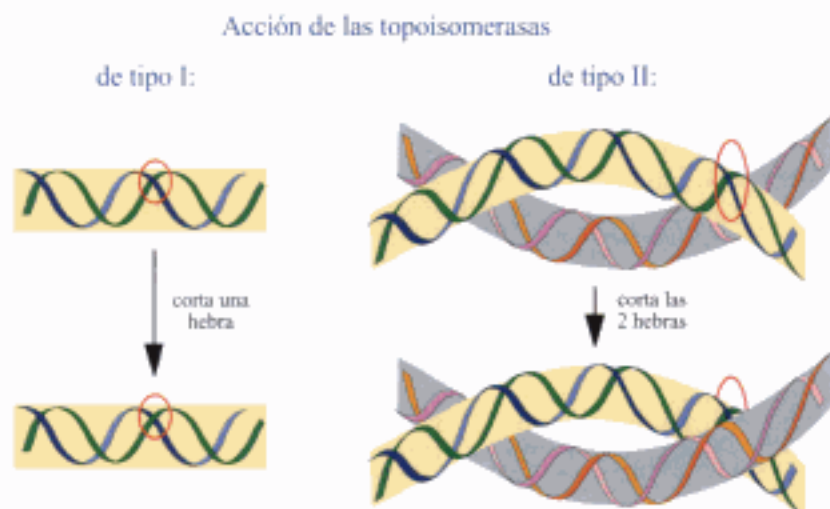
Una vez separadas las hebras por la helicasa, intervienen las "proteínas de unión al DNA de hebra sencilla", que evitan el reapareamiento de las hebras. De esta forma, las bases nitrogenadas pueden acoger a los nucleótidos que se incorporan en la replicación. Se las denomina comúnmente **proteínas SSB** (*single strand binding proteins*), en especial en procariotas. En eucariotas desempeña esta función la **proteína de replicación A (RPA)**.

c) Topoisomerasas

La separación de hebras ocasionada por la progresión de la helicasa a lo largo de la cadena de DNA conlleva la aparición de superenrollamientos positivos por delante de la horquilla. Si se tratase de un DNA lineal y corto, el superenrollamiento se podría eliminar por simple rotación de la hélice en sentido contrario. Sin embargo, ésta no es posible en los DNAs circulares (procariotas) ni en los lineales de gran tamaño (eucariotas), que resultan forzados topológicamente, debido a su naturaleza circular y a su unión a proteínas, respectivamente. Para aliviar esa tensión de torsión se requiere algún tipo de mecanismo giratorio; de lo contrario, el superenrollamiento acumulado por delante de la horquilla llegaría a impedir el avance de la replicación. Este problema se resuelve gracias a la acción de las topoisomerasas, enzimas que alteran el estado de superenrollamiento del DNA sin modificar en otros aspectos su estructura. Es decir, interconvierten topoisómeros modificando el número de enlace del DNA (n^2 de vueltas que da una hebra respecto a la otra, pág. 68). Las topoisomerasas también intervienen en los procesos de recombinación y, posiblemente, transposición, constituyendo un mecanismo básico general de modificación topológica de la molécula de DNA.

Atendiendo a su modo de acción, se distinguen dos grupos de topoisomerasas:

Las **topoisomerasas de tipo I** escinden transitoriamente una hebra del DNA (un enlace fosfodiéster), dejan pasar la otra por la brecha abierta y vuelven a soldar o ligar la primera hebra. Como consecuencia, el número de enlace aumenta o disminuye en una unidad. La reacción transcurre a través de un intermediario covalente entre uno de los extremos del corte y un residuo de tirosina de la enzima. Al permanecer la enzima unida así al DNA abierto durante la acción catalítica, se controla en todo momento el grado de tensión liberada.



Las **topoisomerasas de tipo II**, cuyo ejemplo clásico es la **girasa** procariótica, cortan transitoriamente las dos hebras del DNA, permiten el paso de otra doble hebra, continua, por la brecha abierta y finalizan resellando los cortes. Por consiguiente, el número de enlace aumenta o disminuye en dos unidades. La misión de estas enzimas es generalmente inducir superenrollamientos negativos que facilitan el posterior desapareamiento local. El consumo energético del proceso es importante: se hidroliza una molécula de ATP por cada corte de la cadena.

d) Primasa

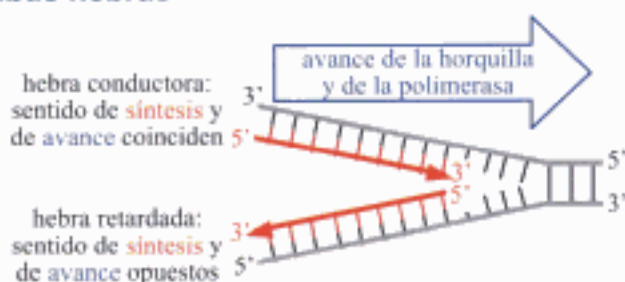
Como ya se ha indicado, las DNA polimerasas no son capaces de iniciar la síntesis de una molécula de DNA, es decir, unir los dos primeros nucleótidos, sino sólo de elongar una hebra preexistente; por ello, el inicio de las hebras nuevas requiere un cebador. Éste es sintetizado por una polimerasa de RNA dirigida por DNA (pág. 147), denominada **primasa** (de *primer*, cebador en inglés), que produce cadenas cortas de RNA, complementarias de cada una de las hebras desparejadas en el origen de replicación. En procariotas, la primasa es una proteína independiente que forma parte de un complejo proteico denominado *primosoma*, asociado a la horquilla. En eucariotas, la actividad primasa reside en una de las subunidades de la DNA polimerasa α , distinta de la subunidad con el centro activo polimerasa (tabla en pág. 149); por ello, se habla de **DNApol α /prim**.

12.3.2 Elongación

La replicación depende no sólo de la DNA polimerasa, sino de numerosas enzimas y factores proteicos (más de 20 en *E. coli*). Este conjunto de proteínas que, como se ha indicado, recibe el nombre de *replisoma* o *complejo de replicación*, se asocia en torno a la horquilla de replicación, para participar de forma directa o indirecta en la síntesis simultánea de ambas hebras.

12.3.2.1 Asimetría de la replicación en ambas hebras

Conceptualmente, el principal problema de la replicación surge al constatar que ambas hebras son copiadas simultáneamente, a medida que avanza la horquilla, mediante enzimas (DNA polimerasas) que sólo son capaces de sintetizar en dirección $5' \rightarrow 3'$. Al abrirse la horquilla, una de las hebras progenitoras se va exponiendo en el sentido correcto para actuar de molde, $3' \rightarrow 5'$, por lo que la síntesis de su hebra complementaria puede transcurrir por simple avance de la polimerasa a lo largo del molde a la par que avanza la horquilla. A esta hebra nueva se la llama **conductora**, **líder** o **guía** (*leading strand*); en ocasiones, su hebra molde recibe el mismo nombre, pero para evitar confusiones la llamaremos "**molde de la hebra conductora**". Por el contrario, la otra hebra progenitora se expone en sentido $5' \rightarrow 3'$, por lo que no puede actuar como molde a medida que avanza la



la otra hebra progenitora se expone en sentido $5' \rightarrow 3'$, por lo que no puede actuar como molde a medida que avanza la

horquilla. Como se verá en seguida, la síntesis de su complementaria requiere un mecanismo particular que, entre otras cosas, supone un desfase con la síntesis de la hebra conductora, por lo que se la llama hebra **retardada** o **retrasada** (*lagging strand*).

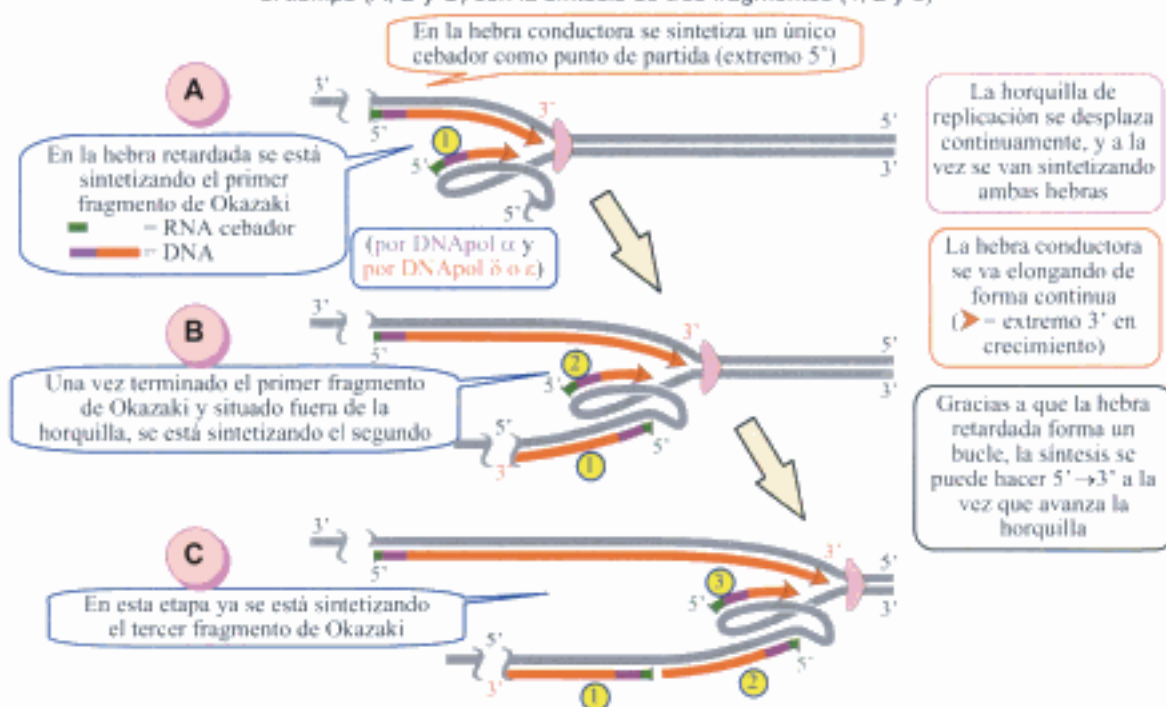
La solución al problema de síntesis de la hebra retardada se reveló al observar que durante la replicación era posible aislar moléculas de pequeño tamaño de DNA recién sintetizado (menor en eucariotas que en procariotas), conocidas como **fragmentos de Okazaki** (por Reiji y Tuneko Okazaki, que los descubrieron). Se propuso entonces un mecanismo que explicase la síntesis de la hebra retardada en paralelo al desplazamiento de la horquilla de replicación, consistente en la síntesis discontinua de la hebra retardada: cada fragmento de Okazaki corresponde a una porción de dicha hebra, cuya síntesis sólo comienza (de ahí el nombre) cuando el avance de la horquilla ha liberado suficiente longitud de hebra sencilla como para que la polimerasa la utilice como molde, en sentido $3' \rightarrow 5'$. Se propone para ello que la hebra sencilla forma un bucle, arrollado sobre la propia enzima, para permitir la síntesis en el sentido $5' \rightarrow 3'$ propio de la polimerasa. La síntesis conjunta de ambas hebras (la conductora, de forma continua, y la retardada, de forma discontinua), por dos moléculas de DNA polimerasa, se dice que es **semidiscontinua**.

Lógicamente, la síntesis de cada fragmento de Okazaki requiere un RNA cebador, sintetizado por una actividad primasa (parte del *primosoma* en procariotas, DNApol α /prim en eucariotas) que viaja con la horquilla y que también sintetizó en el origen de replicación el único cebador requerido por la hebra conductora.

12.3.2.2 Mecanismo de la elongación

Avance de una de las horquillas de replicación:
la síntesis continua de la hebra conductora es simultánea a la
síntesis discontinua de la hebra retardada, en fragmentos de Okazaki

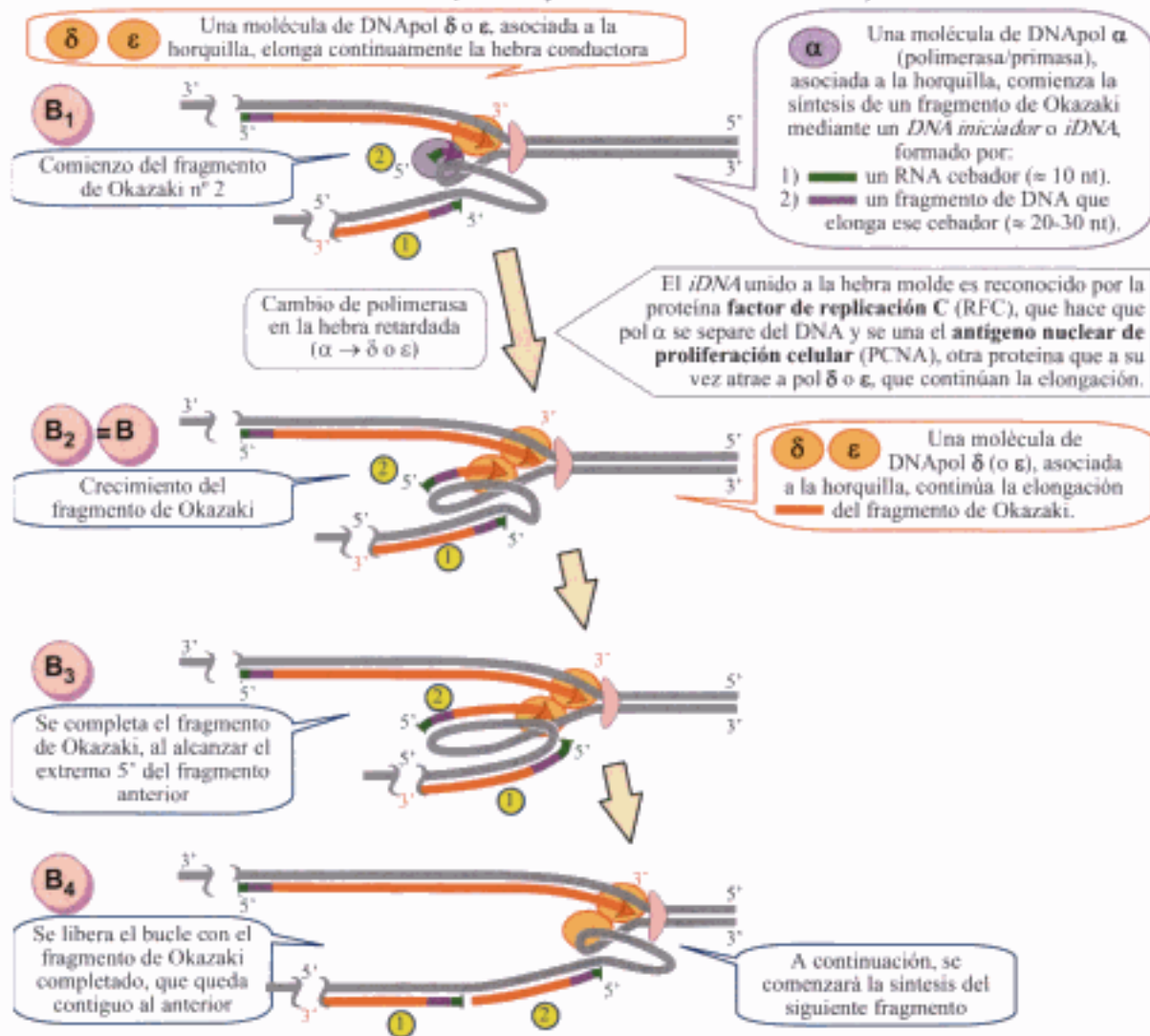
Para una mejor comprensión se muestran 3 etapas sucesivas en el tiempo (A, B y C) con la síntesis de tres fragmentos (1, 2 y 3)



Para que la síntesis de cada fragmento de Okazaki se realice en el entorno de la horquilla, el bucle de hebra molde en torno al replisoma debe sufrir algún tipo de evolución en cuanto a su tamaño y disposición, desde que se comienza la síntesis del fragmento hasta que, una vez terminado, se libera el bucle, ya de doble hebra. Este mecanismo refuerza el concepto de *discontinuidad* en la síntesis de la hebra retardada:

Progreso de la síntesis de un fragmento de Okazaki

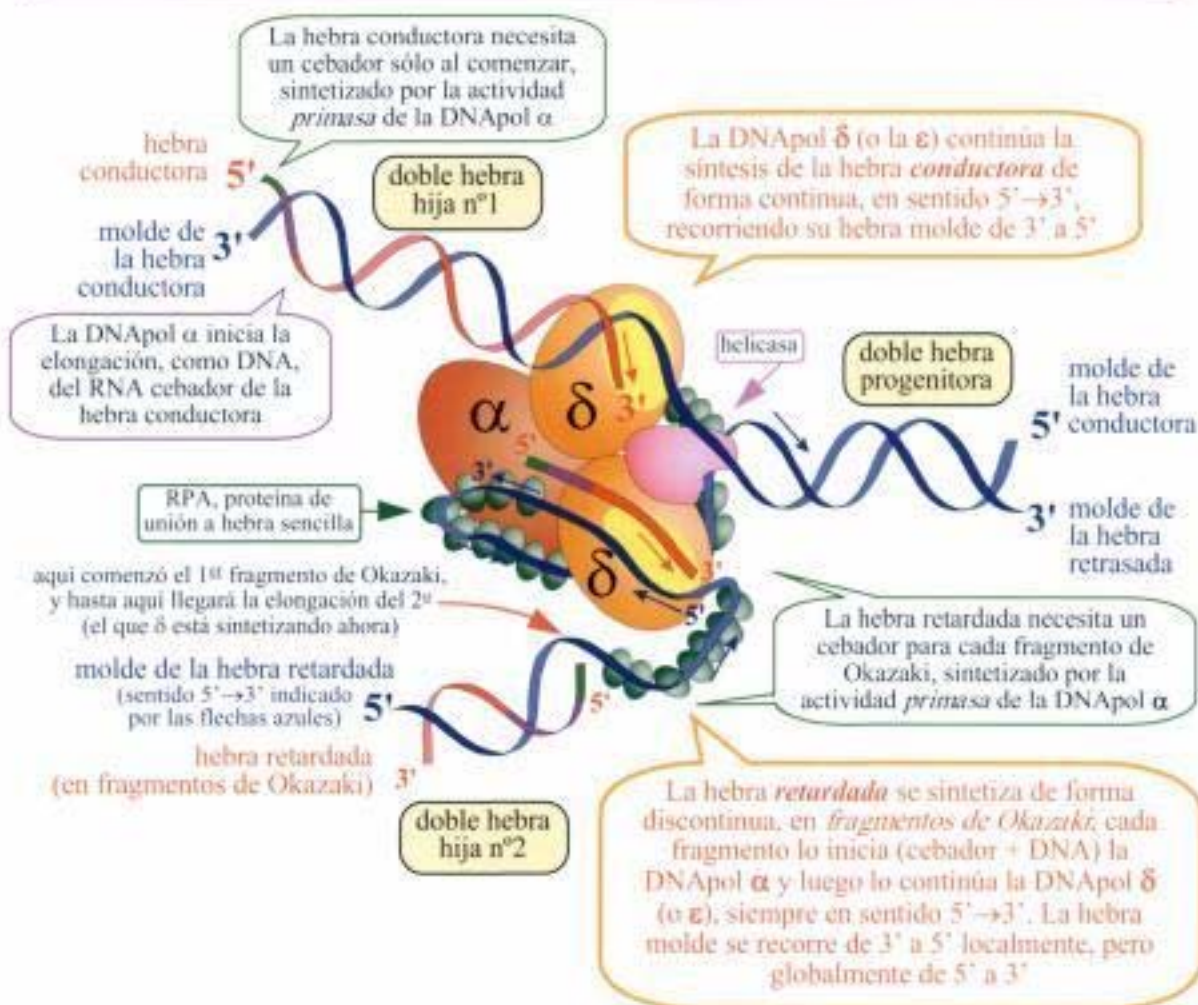
(síntesis del fragmento nº 2, durante el periodo de tiempo necesario para la etapa B de la figura anterior; se incluyen en este caso las DNA polimerasas responsables y la diferente disposición y variación de tamaño del bucle)



Representación global aproximada de la organización de las proteínas responsables de la replicación (componentes del replisoma) en torno a la horquilla de replicación.

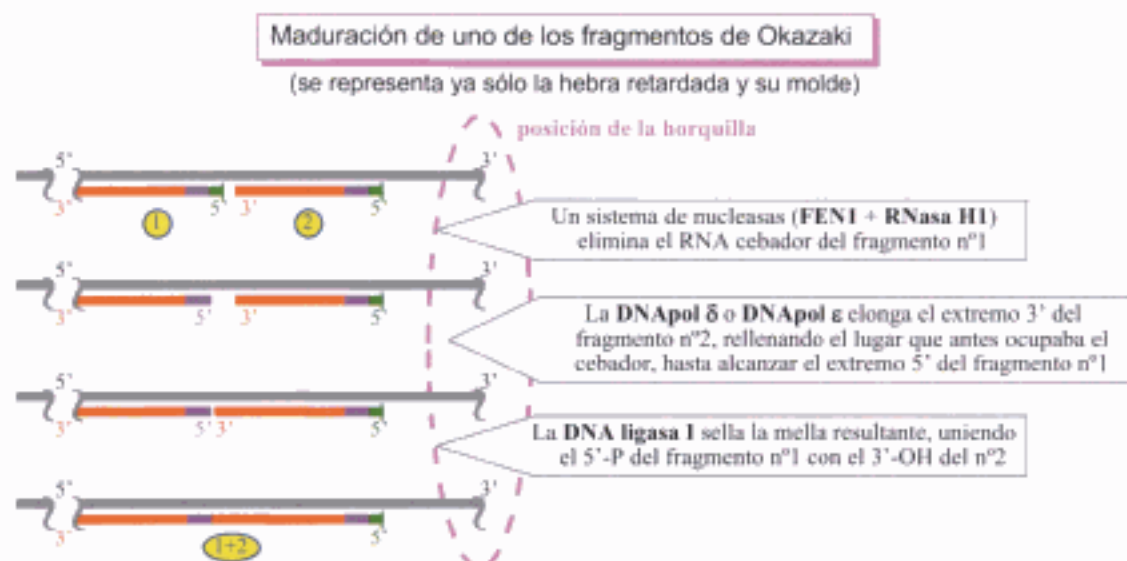
La situación mostrada corresponde a un momento intermedio de la elongación, con un fragmento de Okazaki sintetizado a medias (p.ej., el nº 2 en las figuras anteriores, etapa B).

Por claridad, no se muestran otros factores proteicos, como el PCNA que va asociado a las DNAPol δ , el sistema de nucleasas FEN1/RNasa H1 que madura los fragmentos de Okazaki, la DNA ligasa I, y las topoisomerasas que regulan el superenrollamiento en las dobles hebras progenitora e hijas.

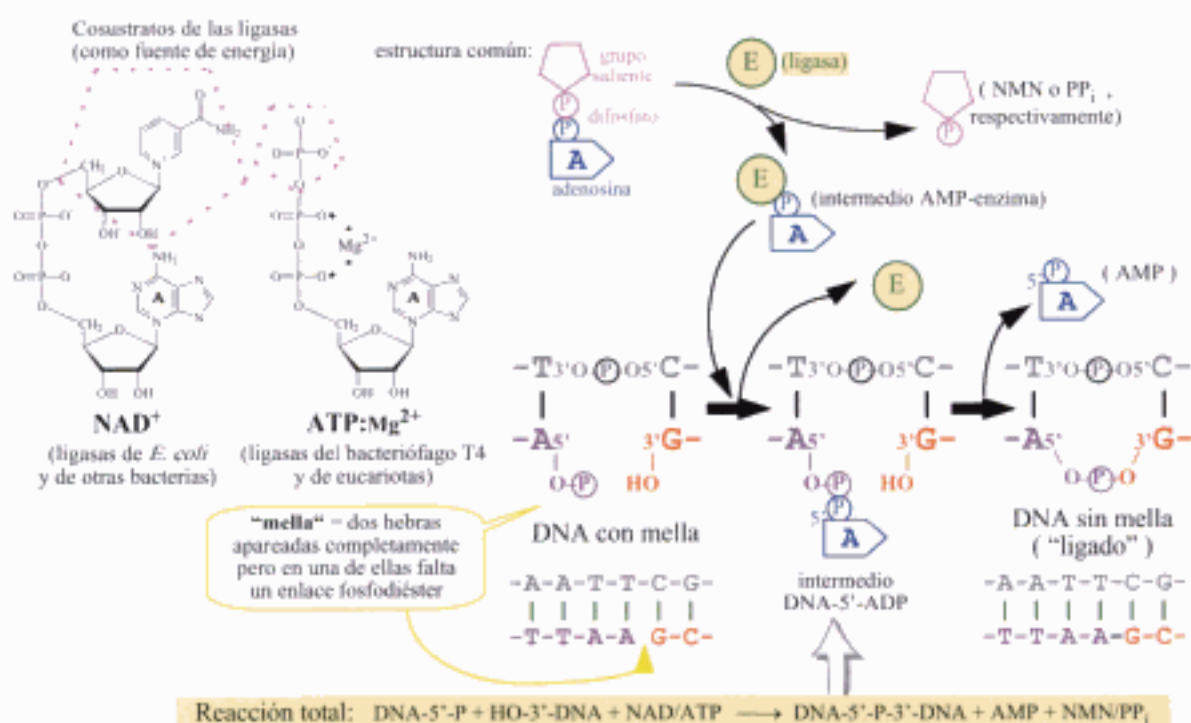


12.3.2.3 Maduración de los fragmentos de Okazaki

Para completar la síntesis de la hebra retardada, deben unirse los fragmentos de Okazaki. Este proceso, denominado globalmente maduración, requiere la eliminación de los cebadores, la elongación del fragmento adyacente para rellenar con DNA el hueco dejado por cada RNA cebador y la unión o empalme de los extremos resultantes para dar una hebra continua. Todo ello tiene lugar posiblemente en las proximidades de la propia horquilla, o bien posteriormente, lejos de ella.



La última etapa en esta maduración de la hebra retardada, la unión de los fragmentos, es catalizada por **DNA ligasas** (o simplemente **ligasas**). El mecanismo químico de la reacción es el siguiente:



El DNA-5'-ADP es la forma activada del extremo 5'-P del DNA que permite el ataque nucleofílico del 3'-OH para regenerar el enlace fosfodiéster

Además de este papel fisiológico, las ligasas son enzimas de gran utilidad en el laboratorio, junto con las nucleasas de restricción, especialmente para la preparación de moléculas de DNA recombinante (tema 16).

12.3.3 Terminación

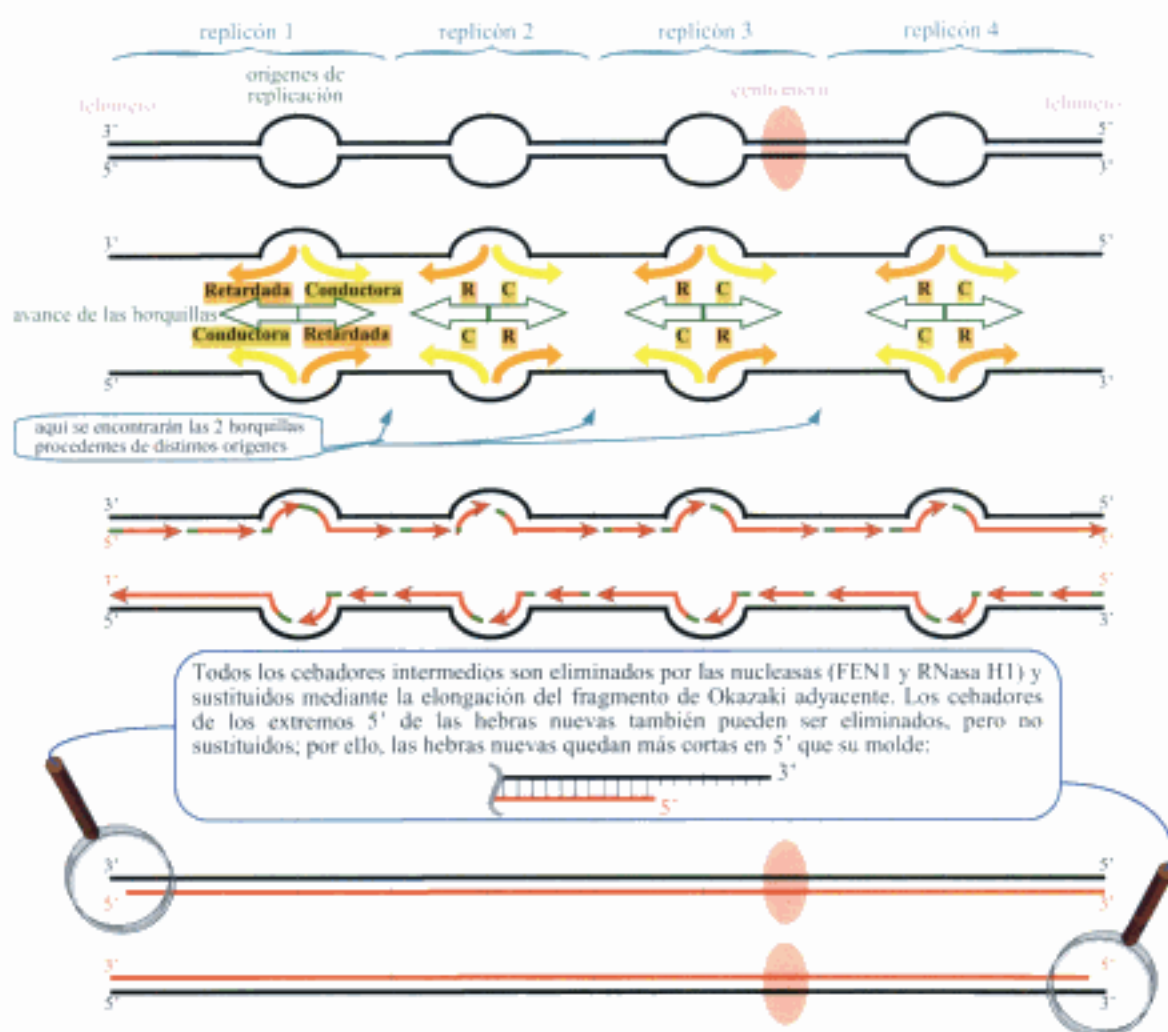
Una vez descritos los detalles particulares del inicio y elongación de la replicación corresponde estudiar los mecanismos de la terminación. Debe indicarse que este proceso no se conoce tan bien como los anteriores, entre otras razones por la propia amplitud y complejidad del concepto de terminación. Se tratará aquí de exponer los problemas que plantea la terminación desde un punto de vista lógico y asumiendo su insuficiente confirmación experimental.

12.3.3.1 Final de la elongación

Se puede considerar, conceptualmente, que la terminación comprende varios aspectos. El primero de ellos es la finalización propiamente dicha de la elongación por la DNA polimerasa, que se verifica en las dos horquillas de replicación de cada replicón, previsiblemente cuando en su avance alcancen a las horquillas respectivas de replicones adyacentes. Éste es un proceso prácticamente desconocido, pero cuya resolución no plantea dificultad conceptual, pues el relleno de huecos y la unión de los dos tramos de nueva hebra puede resolverse de la misma forma que la maduración y unión de los fragmentos de Okazaki en la hebra retardada.

Por otra parte, como segundo problema, falta por aclarar cuál es el número de replicones implicados simultáneamente en la replicación de cada cromosoma, así como los mecanismos de control que coordinan el proceso en todos ellos. A diferencia del caso anterior, no existe aún una hipótesis previsible.

Representación esquemática de la replicación de un cromosoma eucariótico

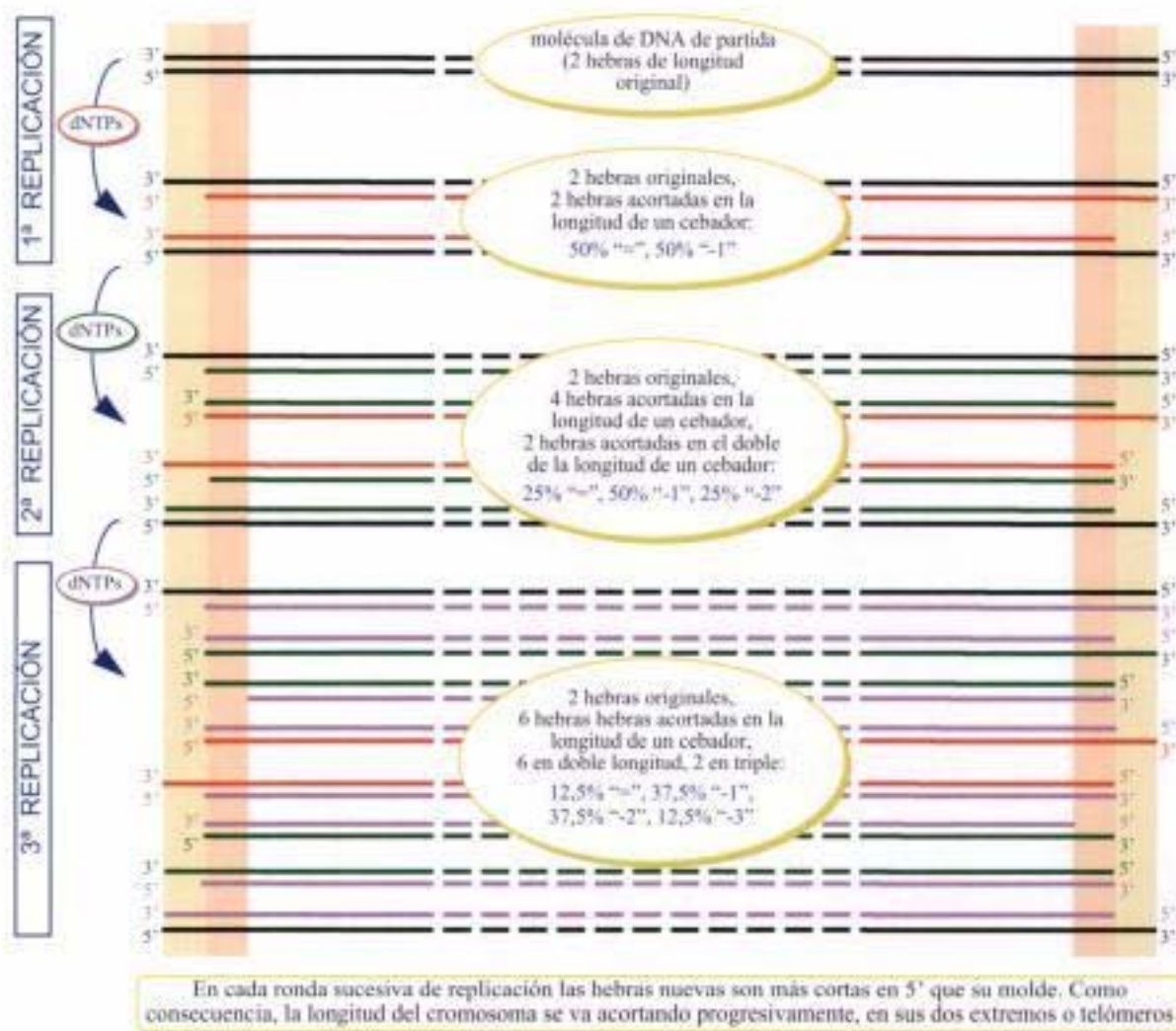


12.3.3.2 Replicación de los telómeros

La replicación de los extremos del cromosoma eucariótico supone el tercer problema de la terminación, pues, como se verá en seguida, no puede llevarse a cabo mediante la elongación por la DNA polimerasa. Este proceso es, por ahora, el mejor conocido de la terminación, posiblemente por su mayor trascendencia fisiológica y patológica.

Como ya se ha estudiado, los telómeros son regiones situadas en los extremos de los cromosomas lineales de la mayoría de eucariotas (pág. 92), a las que se unen ciertas proteínas. Están constituidos por una de las familias de DNA repetitivo no codificante (pág. 113), formando bloques de 1 a 12 kb de longitud, con unidades de 6 a 8 pb repetidas en tándem centenares o miles de veces. Para entender la implicación de los telómeros en la replicación deben recordarse sus dos características principales: la abundancia de repeticiones en tándem y el carácter cohesivo de los extremos (que recuerda, aunque sin relación funcional alguna, a los extremos resultantes de la acción de ciertas enzimas de restricción sobre un DNA, pág. 201).

El carácter cohesivo, en forma de salientes 3' en uno de los extremos de los cromosomas, tanto paternos como filiales, se explica por el mecanismo de replicación discontinua de la hebra retardada (figura anterior). Como consecuencia adicional de dichos salientes 3', se produce un acortamiento progresivo de las hebras de DNA en sucesivas rondas de replicación.

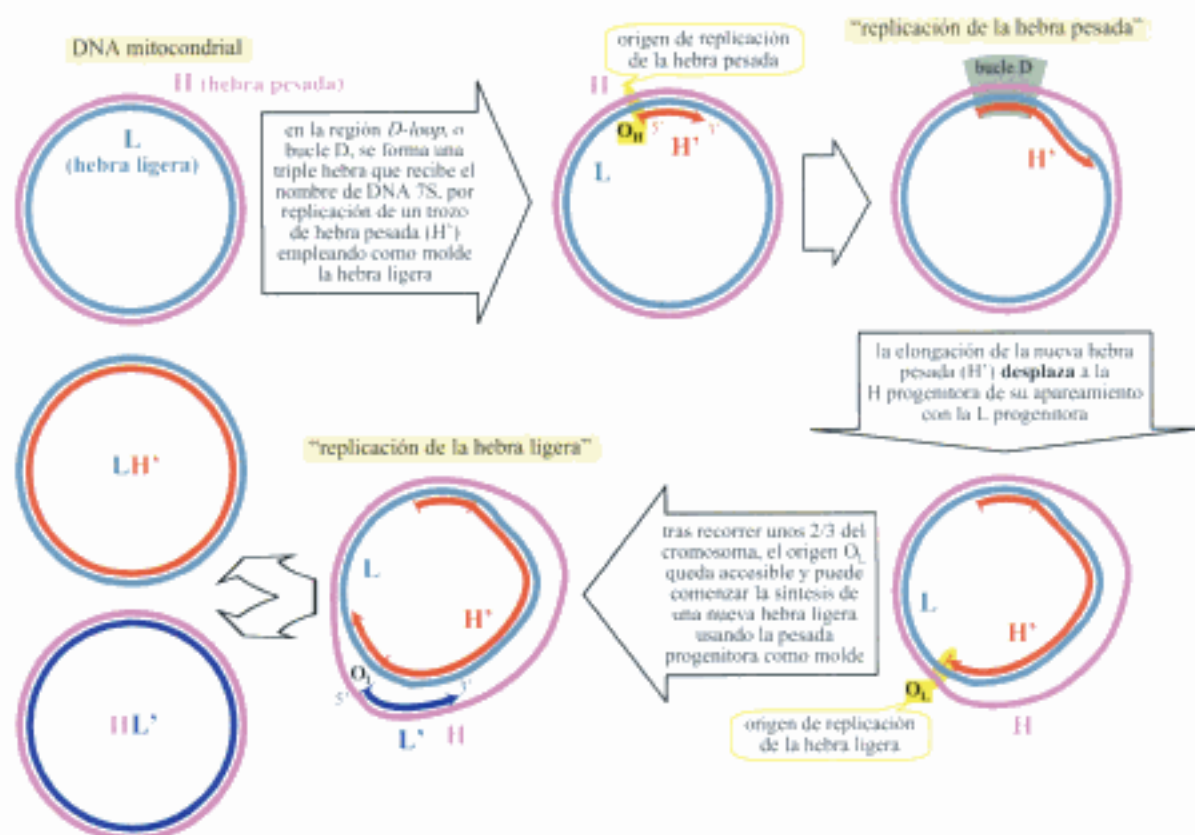


Este acortamiento de las regiones teloméricas (calculado entre 50 y 200 pb por cada división celular) supone la disminución progresiva de la longitud del cromosoma a lo largo de sucesivas divisiones celulares, con el consiguiente problema para la perpetuación del mensaje genético. Este problema se puede resolver en la célula gracias a la actuación, también durante la fase S del ciclo de división celular (tema 9), de un nuevo tipo de enzima, la **telomerasa**, que alarga los telómeros, compensando así la pérdida anterior.

12.3.3.3 Telomerasa: componentes y características

Se ha propuesto que la telomerasa es un resto evolutivo de una ribozima (pág. 76) que pudo servir anteriormente para catalizar la síntesis de DNA; también, que pudo suponer un intermediario en la transición evolutiva desde las RNA replicasas a las transcriptasas inversas proteicas. Se trata de una enzima de tipo DNA polimerasa, concretamente una *polimerasa de DNA dirigida por RNA* (pág. 147) y, por tanto, es una transcriptasa inversa (págs. 196 y 206), pero con características especiales. Efectivamente, se trata de una ribonucleoproteína que sólo es capaz de sintetizar una secuencia determinada de DNA, y que está formada por **dos componentes**:

1. **Componente ribonucleotídico** o "RNA de la telomerasa" (**TR** o **TER**, de *telomerase RNA*), de tamaño variable según especies (desde 146 nucleótidos hasta 1.544), que forma parte integral de la enzima. Incluye una secuencia,



La primera característica singular de la replicación del DNA mitocondrial es que no tiene lugar específicamente en la fase S, sino a lo largo de todo el ciclo celular. Ello está de acuerdo con la división de las mitocondrias durante la interfase, en desfase con la división de otros orgánulos y de la célula. Además, no todas las moléculas de DNA se replican una vez por ciclo, sino que lo hacen al azar.

La segunda peculiaridad que diferencia esta replicación de la del genoma nuclear es su avance **unidireccional**, mediante dos horquillas que parten de orígenes distintos y avanzan en sentido opuesto sintetizando una sola hebra cada una. Como consecuencia de este mecanismo, la replicación también se diferencia de la del DNA nuclear por ser **no simultánea**, **bifocal** y **continua** (no se forman fragmentos de Okazaki, las dos hebras se sintetizan como conductoras).

Carácter	Replicación del genoma nuclear de eucariotas	Replicación del genoma mitocondrial de eucariotas	Replicación del genoma procariótico
Semiconservadora	sí	sí	sí
Simultánea en ambas hebras	sí	no	sí
Secuencial	sí	sí	sí
Bidireccional	sí	no	sí
Origen	múltiples (multifocal)	dos, uno para cada hebra (bifocal)	único (monofocal)
Semidiscontinua	sí	no (continua en ambas hebras)	sí

Tema 13

Hibridación de ácidos nucleicos: fundamento y métodos de ensayo

13.1	Introducción: Desnaturalización y renaturalización del DNA	163
13.1.1	Agentes desnaturalizantes	164
13.1.1.1	Desnaturalización y renaturalización por efecto de la temperatura	164
13.1.1.2	Desnaturalización por ácidos y bases	164
13.1.1.3	Desnaturalización por agentes químicos	165
13.1.2	Influencia de la desnaturalización sobre las propiedades del DNA	165
13.1.2.1	Influencia sobre la viscosidad	165
13.1.2.2	Influencia sobre la absorción ultravioleta	165
13.1.2.3	Significado práctico de las curvas de desnaturalización	166
13.2	Análisis molecular de la hibridación de ácidos nucleicos	167
13.2.1	Principio básico del ensayo de hibridación	168
13.2.2	Influencia de algunos factores sobre la hibridación	168
13.2.2.1	Rigor de la hibridación	168
13.3	Métodos de ensayos de hibridación	169
13.3.1	Hibridación en fase líquida	169
13.3.2	Hibridación en soporte sólido	170
13.3.2.1	Hibridación "dot-blot" y "slot-blot"	170
13.3.2.2	Hibridación de Southern	171
13.3.2.3	Hibridación Northern	171
13.3.3	Hibridación <i>in situ</i>	172
13.3.3.1	Hibridación <i>in situ</i> tisular	172
13.3.3.2	Hibridación <i>in situ</i> cromosómica	172
13.3.4	Técnicas relacionadas	173
13.3.4.1	Transferencia Western	173

La hibridación es un fenómeno de importante aplicación práctica en todas las áreas de la Biología Molecular e Ingeniería Genética, y en especial en la Patología Molecular. Para comprender estas aplicaciones es preciso estudiar previamente la base molecular que fundamenta este proceso y los métodos de ensayo elaborados para su aplicación (en este tema), así como las técnicas de preparación de sondas (tema 14).

La hibridación está basada en el proceso de renaturalización de dos cadenas sencillas de un DNA, que se observa, por ejemplo, al enfriar lentamente una solución del DNA dúplex previamente desnaturalizado por calor. La sensibilidad del apareamiento por puentes de hidrógeno entre bases complementarias (A-T, A-U o C-G), presentes en cadenas sencillas contiguas, da lugar a estructuras de doble hebra, de gran estabilidad. Éstas pueden ser **híbridos DNA-DNA**, **híbridos RNA-RNA** (ambos homodúplex) o **híbridos DNA-RNA** (heterodúplex).

13.1 INTRODUCCIÓN: DESNATURALIZACIÓN Y RENATURALIZACIÓN DEL DNA

Puesto que la hibridación surge de la observación del proceso de renaturalización de un DNA previamente desnaturalizado, se estudian como introducción los procesos de desnaturalización y renaturalización.

Al igual que las proteínas, los ácidos nucleicos **se desnaturalizan** por efecto de agentes químicos o físicos, perdiendo su conformación tridimensional, en especial la forma duplohelicoidal del DNA. La desnaturalización se debe tanto a la rotura de los puentes de hidrógeno entre pares de bases como de las interacciones hidrofóbicas entre bases

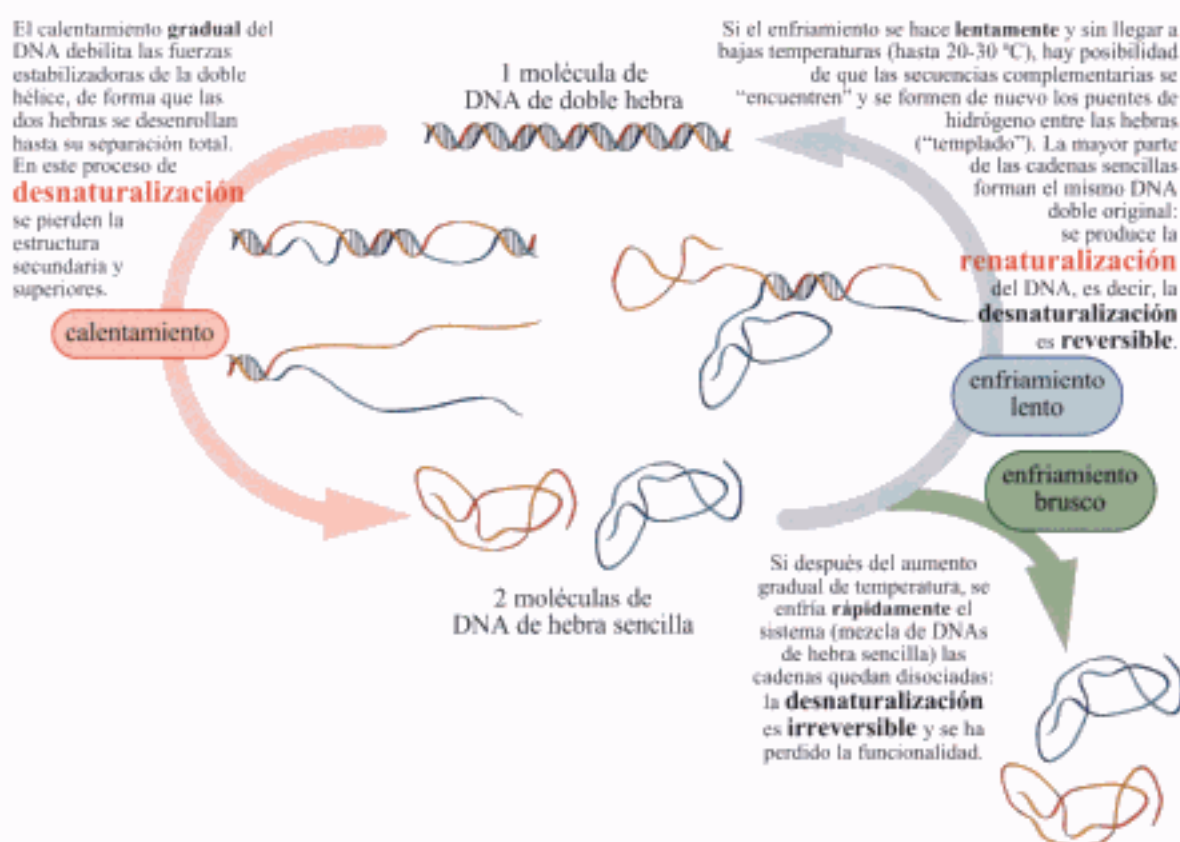
apiladas. Al desnaturalizarse, las dos hebras del DNA se separan y pasan a una conformación al azar (o de ovillo aleatorio), sin que se altere la estructura primaria, pues no hay ruptura de enlaces covalentes. En el caso de los RNAs, con muy distinta conformación tridimensional a pesar de su analogía en estructura primaria, la desnaturalización afecta fundamentalmente a los tRNAs y rRNAs, ya que son los que presentan regiones en doble hélice. Al no dar lugar a aplicaciones experimentales, la desnaturalización del RNA no se considera de forma particular en este tema.

13.1.1 Agentes desnaturalizantes

Aunque pueden considerarse desnaturalizantes los mismos factores que actúan sobre las proteínas, sólo se citarán aquí, brevemente, los ácidos y bases y ciertos agentes químicos, para centrar la atención en el efecto de la temperatura, como factor representativo. Aunque cada uno de ellos tiene un mecanismo particular de actuación, todos debilitan las interacciones por puentes de hidrógeno y/o de apilamiento de bases, conduciendo a un mismo proceso de desnaturalización y de separación de las hebras.

13.1.1.1 Desnaturalización y renaturalización por efecto de la temperatura

La temperatura, el agente desnaturalizante más representativo, actúa de la forma siguiente:



En general, se emplea el proceso de desnaturalización para el estudio de la estructura del DNA, mientras que la renaturalización es útil como base de las técnicas de hibridación, de enorme trascendencia en Biología Molecular e Ingeniería Genética. La desnaturalización/renaturalización es especialmente crítica en la técnica de PCR (tema 15).

13.1.1.2 Desnaturalización por ácidos y bases

Como ya se ha estudiado (pág. 42), el apareamiento de las bases disminuye o se debilita por valores de pH alejados de la neutralidad, con la consiguiente desnaturalización del DNA. Sin embargo, un medio ácido puede conducir también a la ruptura de la estructura primaria de los ácidos nucleicos (recuérdese que el medio ácido fuerte rompe los enlaces fosfodiéster y N-glicosídicos, mientras que en condiciones más suaves –por ejemplo, a pH alrededor de 3– se separan las purinas; pág. 35). Por ello, para la desnaturalización del DNA es más habitual emplear condiciones alcalinas (a las cuales el DNA, a diferencia del RNA, es estable). La neutralización del medio puede permitir la recuperación de la estructura de partida, es decir, la **renaturalización** del DNA (con consideraciones similares a las de la desnaturalización por calor).

13.1.1.3 Desnaturalización por agentes químicos

Algunos de los compuestos empleados con frecuencia para desnaturalizar el DNA son moléculas sencillas, altamente polares, con grupos amino y carbonilo. Éstos les permiten competir en la formación de enlaces de hidrógeno con los grupos amino y carbonilo propios de las bases, facilitando así la ruptura de sus apareamientos y la consiguiente separación de las hebras.



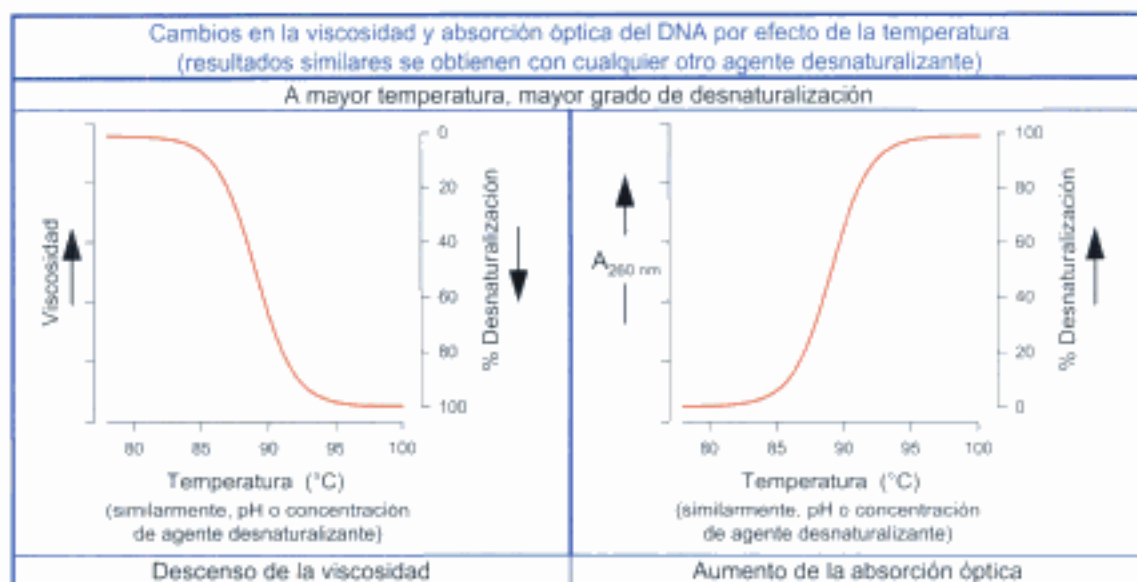
Dependiendo del tipo de acción del agente desnaturalizante y de la forma como ese agente se elimine, se puede conseguir la renaturalización del DNA, de modo análogo a los casos anteriores.

13.1.2 Influencia de la desnaturalización sobre las propiedades del DNA

En la práctica, la desnaturalización del DNA se puede seguir fácilmente midiendo aquellas propiedades físicas que permiten diferenciar el DNA de doble hebra y el de hebra sencilla. Varias de ellas experimentan profundos cambios durante el proceso: aumento de la absorción óptica, cambio en la rotación óptica, aumento de la densidad de flotación y descenso de la viscosidad. Las más utilizadas son la absorción en el ultravioleta (a 260 nm) y la viscosidad. A partir de estos cambios también se puede obtener información de tipo estructural acerca del propio ácido nucleico (aunque este aspecto no se considera aquí).

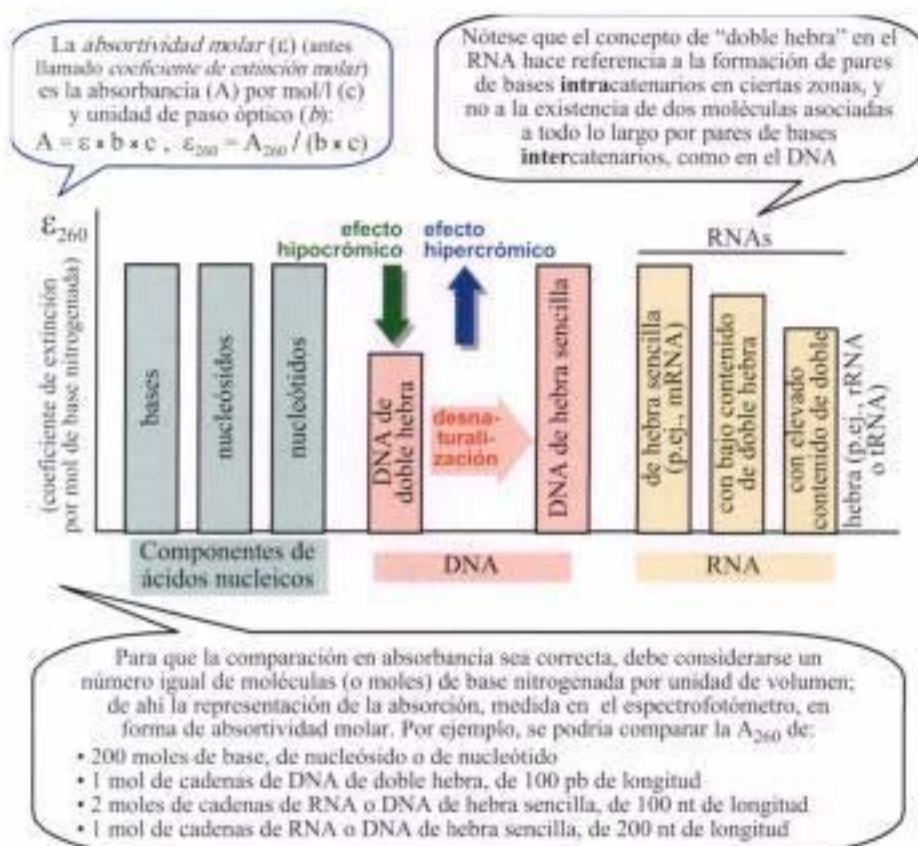
13.1.2.1 Influencia sobre la viscosidad

La relativa rigidez de la doble hélice y su inmensa longitud en relación al diámetro hacen que las disoluciones de DNA de doble hebra sean muy viscosas. Por el contrario, en forma de hebra sencilla el plegamiento al azar de la cadena hace que la viscosidad sea menor. De acuerdo con ello, al aumentar la temperatura, el pH o la concentración del agente químico y producirse la desnaturalización, se observa una disminución en la viscosidad de la disolución de DNA.



13.1.2.2 Influencia sobre la absorción ultravioleta

Recuérdese que las bases nitrogenadas, los nucleósidos y los nucleótidos poseen valores máximos de absorción de la luz cercanos a la longitud de onda de 260 nm, debido a la presencia en su molécula de sistemas conjugados de dobles enlaces. Esta propiedad se mantiene en los ácidos nucleicos, pero la magnitud de la absorción no es la misma, pues el apilamiento y los puentes de hidrógeno la reducen (a esto se le llama **efecto hipocrómico**). Por ello, el DNA de hebra sencilla absorbe en magnitud similar a los nucleótidos libres, mientras que el DNA de doble hebra absorbe menos. Como consecuencia, la desnaturalización del DNA va acompañada de un aumento en la A_{260} (al que se llama **efecto hipercrómico**).



Obsérvese comparativamente la relación existente entre disminución de la viscosidad y aumento de $A_{260 \text{ nm}}$ con el proceso de desnaturalización del DNA.

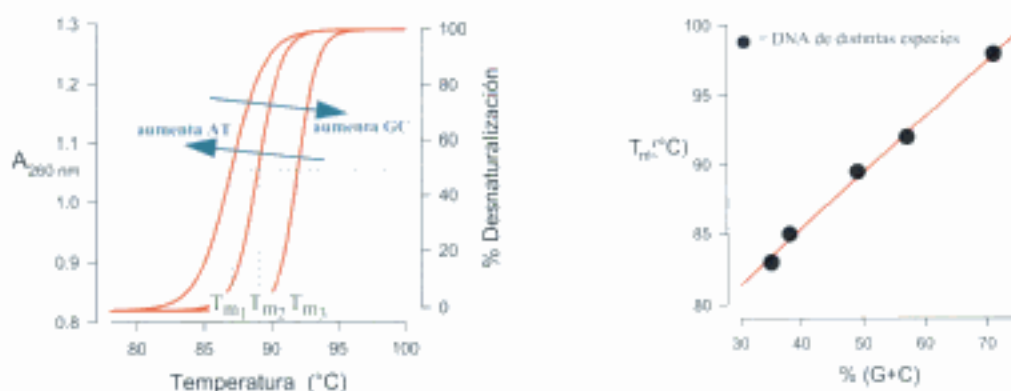
El RNA, al presentar en su molécula un grado variable de doble hebra, presentará una absorción intermedia entre la de un DNA dúplex de tamaño equivalente y uno de hebra simple (desnaturalizado); la magnitud depende de la complejidad de su estructura secundaria y, por tanto, también del tipo de RNA (mensajero, ribosómico o de transferencia).

13.1.2.3 Significado práctico de las curvas de desnaturalización

La representación gráfica de la variación de viscosidad o de A_{260} frente a la temperatura (o concentración de agente desnaturalizante) permite obtener las **curvas de desnaturalización**, que son características para cada DNA en disolución. El carácter sigmoidal de estas curvas indica que la desnaturalización se produce bruscamente, en un margen muy estrecho de temperatura, de pH o de concentración de agente desnaturalizante. Esto contrasta con la transición más gradual observada en la desnaturalización térmica de proteínas, y sugiere que las interacciones que mantienen la doble hélice (apilamiento y puentes de hidrógeno) son cooperativas, es decir, la separación de las hebras en un punto favorece que se sigan separando en el resto de la molécula.

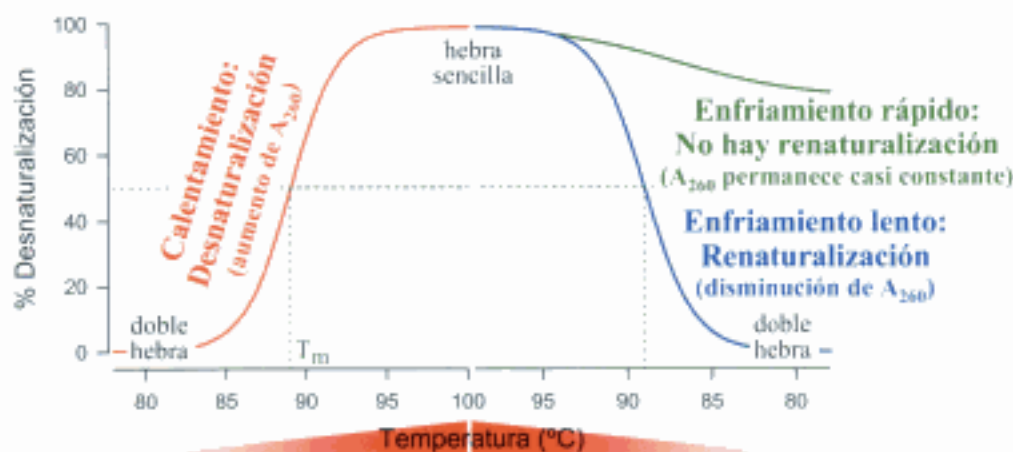
Al ser la transición tan brusca, la desnaturalización térmica del DNA es similar en cierto modo a la fusión de un sólido, por lo que se habla también de **curvas de fusión**, y a la temperatura central del intervalo de transición (50% de desnaturalización) se la llama **temperatura de fusión** (T_f o T_m , del inglés *melting*). Bajo condiciones definidas, cada DNA muestra una curva de fusión con valores característicos de T_m y de magnitud del efecto hipercrómico (diferencia entre A_{260} máxima y mínima, o cambio de viscosidad). Estos datos son una medida de la estabilidad estructural de la molécula de ácido nucleico. Su valor depende además del pH y la fuerza iónica del medio, y de la composición en bases del ácido nucleico.

Es importante resaltar la relación entre T_m y el contenido en pares de bases C=G o A=T del DNA: Cuanto más abundantes son C=G, más desplazada hacia la derecha aparece la curva, es decir, se necesita mayor temperatura para la desnaturalización, T_m es mayor. Y viceversa, cuanto más pares A=T, mayor desplazamiento a la izquierda y menor T_m . Esto se debe a que, al ser más fuerte la interacción entre pares C=G (3 puentes de hidrógeno), necesitan mayor energía para disociarse.



En efecto, al determinar la temperatura de fusión para varios DNAs de orígenes diversos, se ha encontrado una correlación lineal con su contenido de G+C (que, de acuerdo con una de las reglas de Chargaff, ya estudiadas (pág. 39), varía entre especies y puede emplearse como índice en estudios evolutivos). Ello permite utilizar la T_m de una muestra de DNA para estimar su composición en bases. Por ejemplo, se puede utilizar para determinar la presencia de regiones ricas en G+C o de "islotos CpG", asociados frecuentemente en el genoma eucariótico a regiones promotoras, reguladoras de la transcripción y comúnmente metiladas.

Finalmente, la representación gráfica de las curvas de desnaturalización y renaturalización de un DNA permite observar, como se indicó al principio, que la similitud entre ambos procesos depende de las condiciones en las que se intente la renaturalización. Obviamente, ambas curvas discurren en sentidos opuestos, pero sólo si el enfriamiento es lento tiene lugar una renaturalización, mientras que un enfriamiento rápido impide la reasociación correcta de las hebras:



13.2 ANÁLISIS MOLECULAR DE LA HIBRIDACIÓN DE ÁCIDOS NUCLEICOS

Debe tenerse en cuenta que la renaturalización depende sólo de la complementariedad de secuencia de las hebras. Por ello, aunque lo más probable es que éstas se reasocien formando las mismas parejas o dúplex originales, pueden darse otras reasociaciones. En particular, si se parte de una mezcla desnaturalizada de moléculas distintas de DNA, pero con secuencias parecidas en parte (por ejemplo, de DNA que codifica la misma proteína en especies distintas), al renaturalizar se puede producir la asociación parcial de dos hebras procedentes de moléculas dúplex distintas. Se forman así **DNAs híbridos** de doble hebra, con regiones de apareamiento y de no apareamiento. Esta formación de moléculas dúplex de DNA cuyas hebras tienen distinto origen (distinta especie o, en general, distinto ácido nucleico) recibe el nombre de **hibridación**. Es característico el carácter artificial de su consecución, de ahí que también se considere hibridación la formación de híbridos entre DNA y RNA, o entre un ácido nucleico y un oligonucleótido natural o sintético, siempre en función de la existencia de secuencias total o parcialmente complementarias.

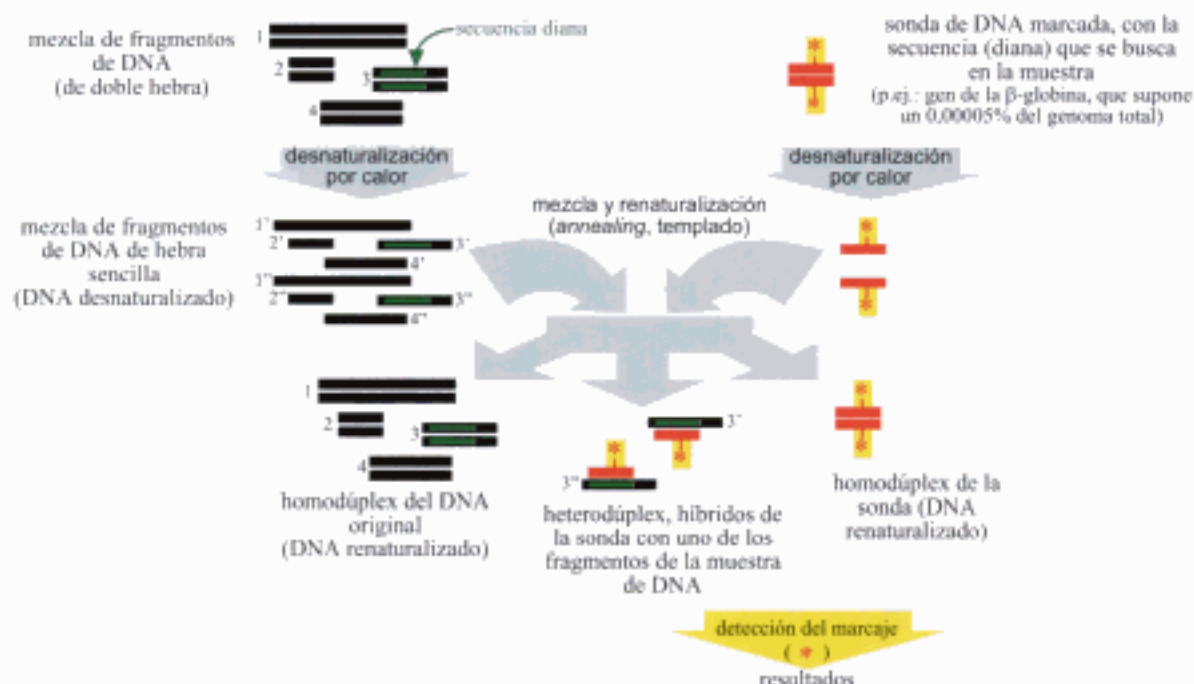
La estabilidad del híbrido es tanto mayor cuanto más elevada sea la proporción de bases complementarias y mayor la longitud de las secuencias apareadas. El hecho de que cadenas diferentes de ácido nucleico hibriden entre sí indica que los organismos de donde proceden comparten un cierto grado de herencia evolutiva común, y que sus RNAs y proteínas tienen estructuras y funciones similares. Se dice que los ácidos nucleicos que codifican estas proteínas y RNAs tienen **secuencias** similares, **homólogas**. A mayor grado de hibridación, mayor conexión evolutiva entre especies, y viceversa; por ejemplo, el DNA humano hibrida en mayor proporción con el DNA de ratón que con el de

levadura. Sin embargo, no es ésta la única utilidad de la hibridación, sino que da lugar a numerosas aplicaciones, muchas entre ellas en el diagnóstico clínico.

13.2.1 Principio básico del ensayo de hibridación

Todos los ensayos de hibridación se basan en la gran especificidad de la interacción entre bases complementarias. Para que la hibridación permita identificar en un genoma, u otra muestra cualquiera de ácido nucleico, una secuencia particular se requieren dos **elementos básicos**:

- La presencia de la **secuencia diana** en uno de los fragmentos, generalmente obtenidos con enzimas de restricción, de la **muestra** de ácido nucleico.
- Un fragmento corto de ácido nucleico, llamado **sonda**, de secuencia conocida y complementaria a la secuencia diana, marcado de forma que permita su detección (radiactivo, coloreado, fluorescente...; tema 14).



13.2.2 Influencia de algunos factores sobre la hibridación

Además de la composición de bases de las cadenas, existe una serie de parámetros que afectan a la posibilidad de renaturalización de dos hebras sencillas y, por tanto, condicionan experimentalmente la realización de un ensayo de hibridación. Estos factores, cuya descripción detallada corresponde a un tratado específico de metodología en Biología Molecular, están esencialmente relacionados con la concentración de DNA o RNA diana (no sólo la concentración de ácido nucleico en la muestra, sino el número de copias de la secuencia diana que contiene esa muestra), la concentración y tamaño de la sonda, la temperatura y tiempo de hibridación, y las características del medio de reacción, tales como fuerza iónica, pH, concentración de agentes desnaturalizantes, adición de ciertos polímeros, etc. A continuación se comenta sólo un aspecto, de interés general en cualquier ensayo de hibridación, determinado por algunos de dichos factores.

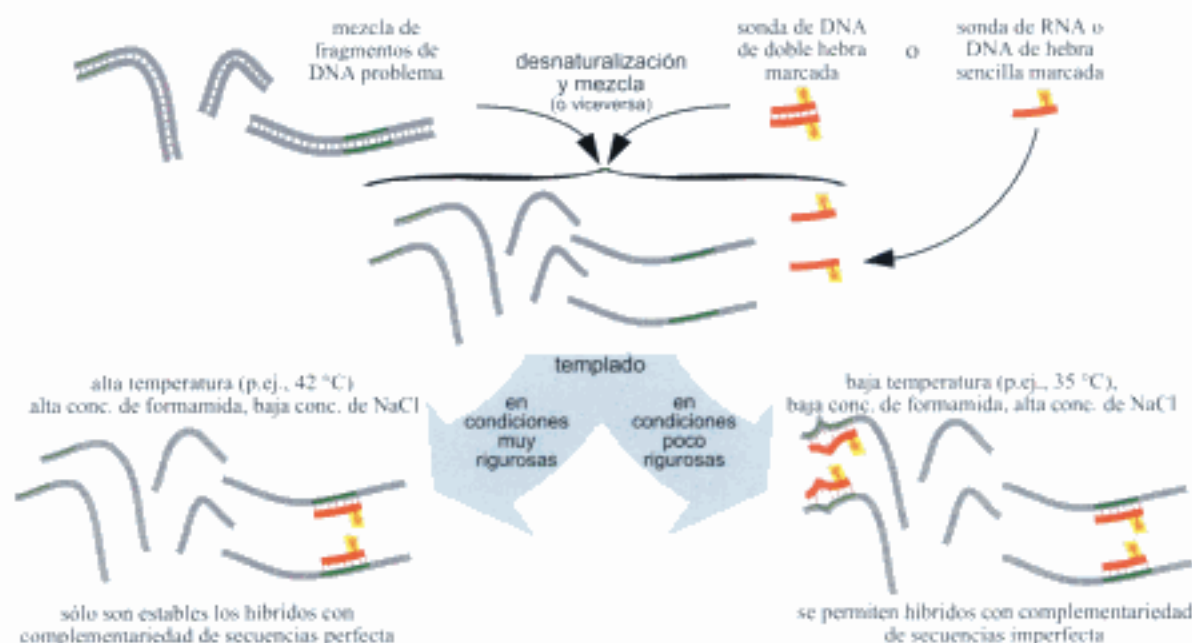
13.2.2.1 Rigor de la hibridación

Se define el **rigor** (en inglés, *stringency*) como el grado de especificidad en el apareamiento conseguido en los híbridos; es decir, supone una medida de la capacidad de reconocimiento mutuo de dos secuencias. Al realizar la hibridación, normalmente se busca un apareamiento completo entre sonda y diana, pero en ocasiones interesa tolerar un cierto grado de desapareamiento, con el fin de detectar regiones del DNA cuya secuencia no es exactamente igual a la diana, pero presenta elevada homología. Por otro lado, el control del rigor en cada aplicación y caso concreto es esencial: ni debe ser tan alto que impida la unión estable de la sonda a su diana, ni tan bajo que permita la hibridación inespecífica de la sonda con secuencias poco relacionadas, dando lugar a una señal de fondo elevada o a falsos positivos.

El rigor depende de diversos factores experimentales: temperatura y tiempo de renaturalización (o templado), longitud y concentración de las cadenas, composición de bases y ambiente químico (presencia de cationes, formamida, urea, pH, tipo de disolución, etc.). Cuanto más elevada es la temperatura durante la fase de templado, más específica

debe ser la interacción entre las hebras para que se mantengan unidas. Los cationes monovalentes (Na^+) estabilizan el dúplex, pues apantallan la repulsión electrostática entre los grupos fosfato, facilitando la aproximación de ambas hebras. Por último, la formamida y la urea afectan a los enlaces de hidrógeno (pág. 165) y son por ello agentes que desestabilizan la doble hélice.

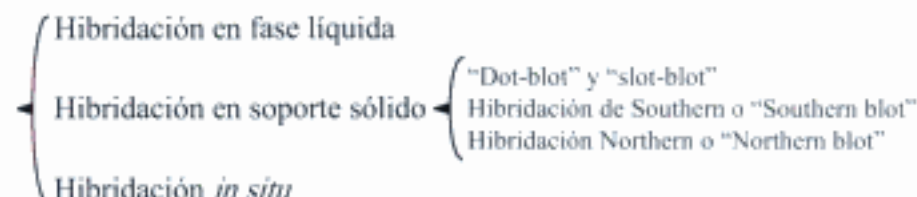
En consecuencia, los principales factores que se suelen manejar en la práctica para maximizar el rigor de la hibridación, y permitir sólo los híbridos que estén apareados completamente, son la elevación de la temperatura de templado, la disminución de la concentración de Na^+ y el aumento de la de formamida. Viceversa, para reducir el rigor de la hibridación, es decir, para permitir la formación de híbridos parcialmente despareados, se emplean temperaturas de templado relativamente bajas, concentraciones altas de Na^+ y bajas de formamida.



13.3 MÉTODOS DE ENSAYOS DE HIBRIDACIÓN

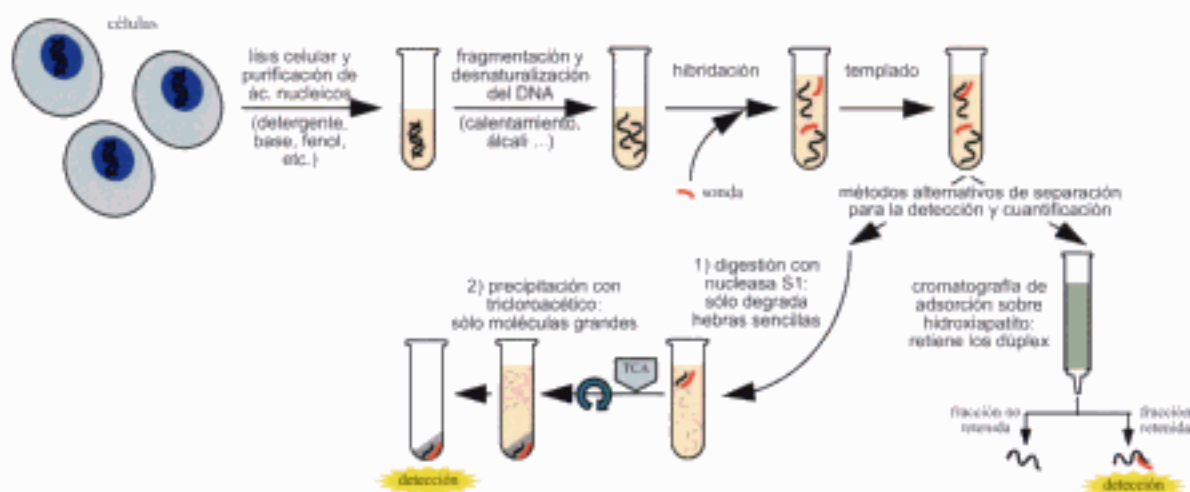
Todos los ensayos de hibridación se basan en la mezcla de hebras sencillas de ácido nucleico muestra o diana, no marcado, con una sonda de secuencia conocida, marcada, bajo condiciones experimentales que permitan el apareamiento entre bases complementarias. Si la sonda o el DNA diana son de doble hebra, han de ser previamente desnaturalizados, generalmente por calor o por tratamiento alcalino. Una vez mezcladas, se permite la **renaturalización** en la cual, como ya se ha comentado, se pueden formar híbridos sonda:diana. Se ha de disponer, además, de un método para detectar los híbridos formados, que viene determinado por el tipo de marcaje empleado en la sonda.

Los ensayos de hibridación se pueden clasificar en tres grupos:



13.3.1 Hibridación en fase líquida

Éste fue el método inicialmente empleado. Utilizando una muestra de DNA o RNA en disolución, se realiza su hibridación con la sonda; al estar ambas disueltas, la cinética es más rápida. Además, todas las moléculas diana son igualmente accesibles, por lo que la técnica es más cuantitativa que la hibridación en soporte sólido. Sin embargo, es el método que menos se utiliza actualmente, habiendo sido desplazado por otros más convenientes. Se aplica, por ejemplo, en el estudio de la estructura de los mRNAs empleando nucleasa S1.

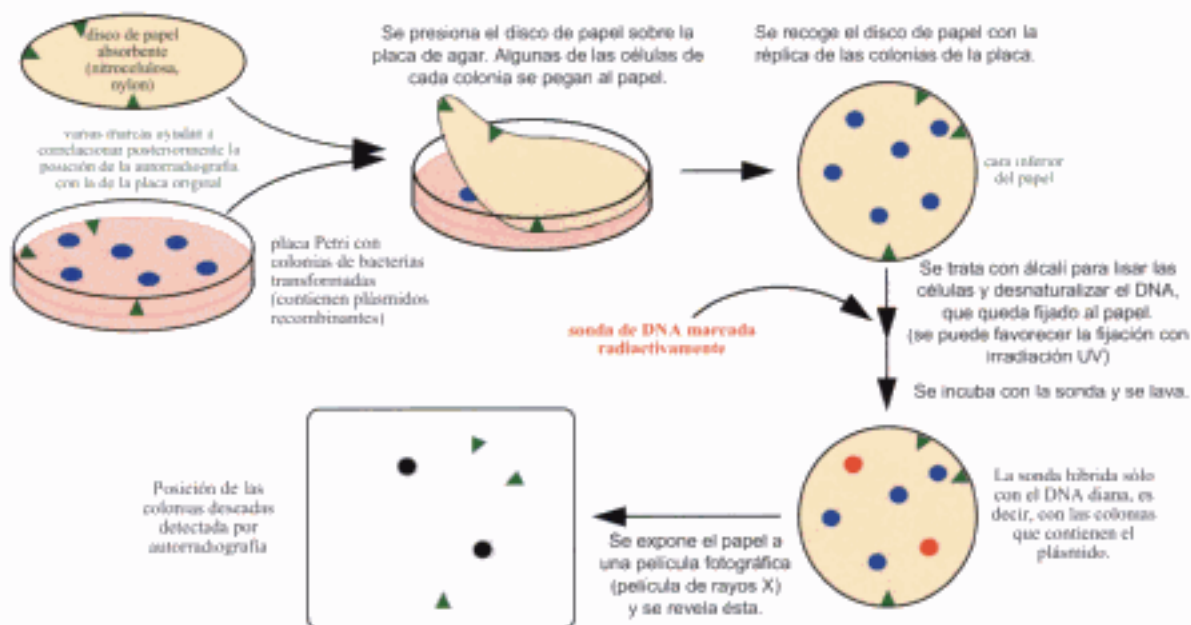


13.3.2 Hibridación en soporte sólido

Se trata de un tipo de ensayo más simple, que permite procesar varias muestras simultáneamente y facilita el control de la hibridación. Aunque la cinética de la hibridación es más lenta y el proceso menos eficiente, se utiliza mucho más que la hibridación en disolución, por ser más sencillo de realizar y más versátil.

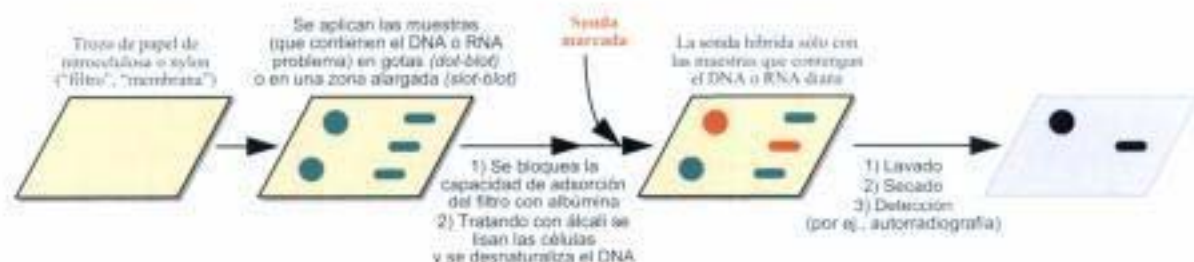
Estos métodos han adquirido hoy día un interés particular como herramientas muy válidas para el análisis de ácidos nucleicos y el diagnóstico de enfermedades moleculares con base genética. Como antecedentes, pueden citarse la inmovilización del DNA en nitrocelulosa, su detección por primera vez por hibridación y el descubrimiento de un método para la lisis *in situ* de colonias bacterianas en filtros de celulosa.

El **esquema general** de este tipo de ensayos se puede ilustrar con el método comúnmente empleado para identificar, de entre varias colonias bacterianas en placas de cultivo con agar, aquellas que contienen un gen o fragmento de DNA de interés bajo la forma de plásmido recombinante (tema 16).



13.3.2.1 Hibridación "dot-blot" y "slot-blot"

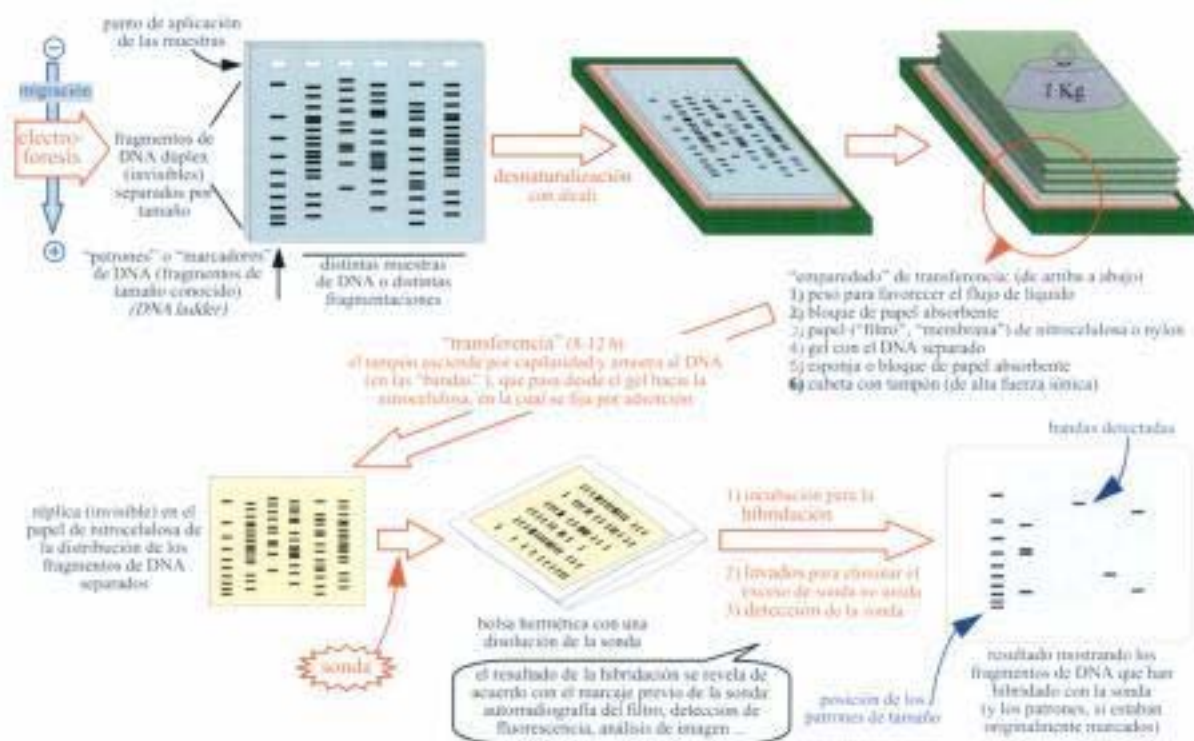
Es el método más simple y común, gracias a su posibilidad de empleo con muestras de ácido nucleico sin purificar (por ejemplo, sangre, heces o esputos). La muestra se aplica, bien gota a gota (*dot-blot*) o en manchas alargadas (*slot-blot*), sobre un "filtro" de nylon o nitrocelulosa, cuyas propiedades de adsorción fijan las moléculas de ácido nucleico y proteína.



13.3.2.2 Hibridación de Southern

Desarrollada por E.M. Southern, al que debe su nombre (*Southern blot*). Es una técnica simple y fácil de realizar que detecta fragmentos de DNA, separados por tamaño mediante electroforesis en gel (pág. 137). Combina la separación electroforética del DNA con su transferencia a un soporte sólido o "filtro" (nylon o nitrocelulosa) para su hibridación. Tanto ésta como la posterior detección se realizan de forma más eficaz en el filtro de lo que se haría en el gel. Esta técnica ha contribuido de forma notable a aumentar la potencialidad de la hibridación como herramienta analítica, constituyendo hoy día en bioquímica clínica el método habitual de análisis del DNA extraído de muestras de sangre, esputos, tejido y células (linfocitos, amniocitos...).

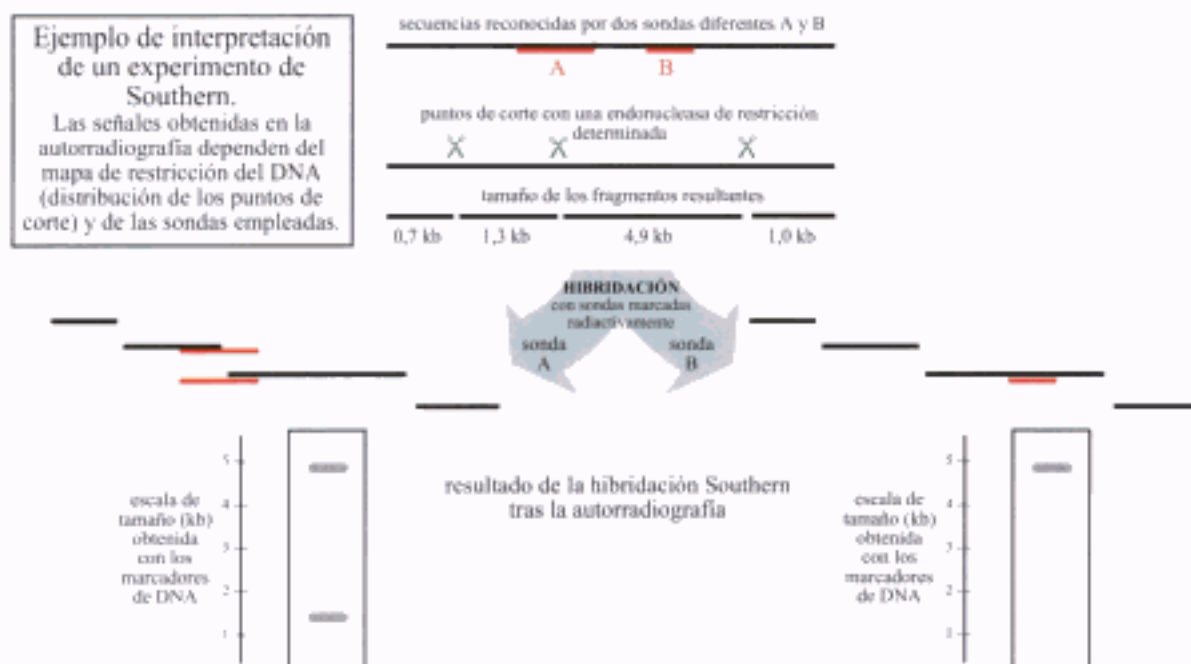
Metodológicamente, implica cinco etapas bien diferenciadas: separación electroforética, desnaturalización, transferencia, hibridación y detección:



13.3.2.3 Hibridación Northern

A diferencia del caso anterior, su nombre (*northern blot*) no corresponde al de ningún investigador, sino a un simple juego de palabras (*southern* = del sur, *northern* = del norte). El procedimiento es idéntico, pero el ácido nucleico analizado es RNA. Por su menor tamaño, no es necesaria su fragmentación con endonucleasas, pero debe tenerse especial cuidado para evitar la acción de RNasas, tanto endógenas como exógenas (contaminación muy común, incluso por las manos del analista, y muy difíciles de inactivar). El uso de agentes desnaturalizantes sigue siendo necesario para evitar apareamientos intracatenarios, que alterarían el patrón de separación por tamaño.

La técnica de hibridación Northern se emplea principalmente para obtener información sobre el tamaño de RNAs y sobre el modelo de expresión de genes específicos. Éstos, una vez clonados, pueden ser utilizados como sondas para hibridar muestras de RNA aisladas de una variedad de tejidos diferentes, y así conocer los tipos celulares en los que se expresa el gen, la abundancia relativa y tamaño de los transcritos, y diversos detalles de su procesamiento posterior.



13.3.3 Hibridación *in situ*

Es un tipo especializado de hibridación en soporte sólido, en la que la sonda se aplica sobre muestras en su ubicación natural, generalmente aquéllas preparadas para microscopía. Supone, por tanto, el empleo de la hibridación bajo un abordaje citológico e histológico, y metodológicamente es análoga a las técnicas inmunocitoquímicas e inmunohistoquímicas. De forma muy global, cabe distinguir entre sus aplicaciones tisular y cromosómica.

13.3.3.1 Hibridación *in situ* tisular

Se realiza sobre cortes de tejido, células intactas, núcleos aislados... En muchos casos, se lleva a cabo directamente sobre tejidos fijados con formalina, incluidos en parafina, congelados, etc. Es una modalidad de hibridación particularmente importante para detectar virus patógenos, para diagnosticar células cancerosas como consecuencia de cambios genéticos o para estudiar cambios en la expresión génica, detectando secuencias en el RNA celular. En este caso, tras una fijación suave de la preparación, se añaden sondas de DNA complementario (cDNA), de hebra sencilla. La detección se realiza por autorradiografía y examen de la película fotográfica bajo el microscopio, o por microscopía de fluorescencia, si se empleó este marcaje para la sonda (tema 14).

13.3.3.2 Hibridación *in situ* cromosómica

Se realiza sobre preparaciones en porta de cromosomas metafásicos o prometáfásicos (pág. 90), tratadas con fijadores para preservar las características morfológicas del cromosoma. Para hacer accesible la secuencia de DNA, se trata la muestra con enzimas proteolíticas y se desnaturaliza el DNA por calor, álcali o formamida para permitir su hibridación con la sonda. Inicialmente se emplearon sondas radiactivas, pero problemas técnicos en la detección de la señal limitan su utilidad. Se ha conseguido aumentar la sensibilidad y resolución con el empleo de sondas marcadas con fluorescencia. En esta técnica, llamada **hibridación *in situ* por fluorescencia** o **FISH** (*fluorescence in situ hybridization*), la detección se hace bajo microscopio de fluorescencia, bien de forma directa o por un método indirecto (pág. 185).

En cuanto a sus aplicaciones, dada la gran detectabilidad de los fluorocromos actuales se puede emplear la FISH con fines muy diversos y se ha convertido en una técnica con elevado potencial diagnóstico en cromosopatías (tema 29), genes tumorales, virología, trasplantes, etc.

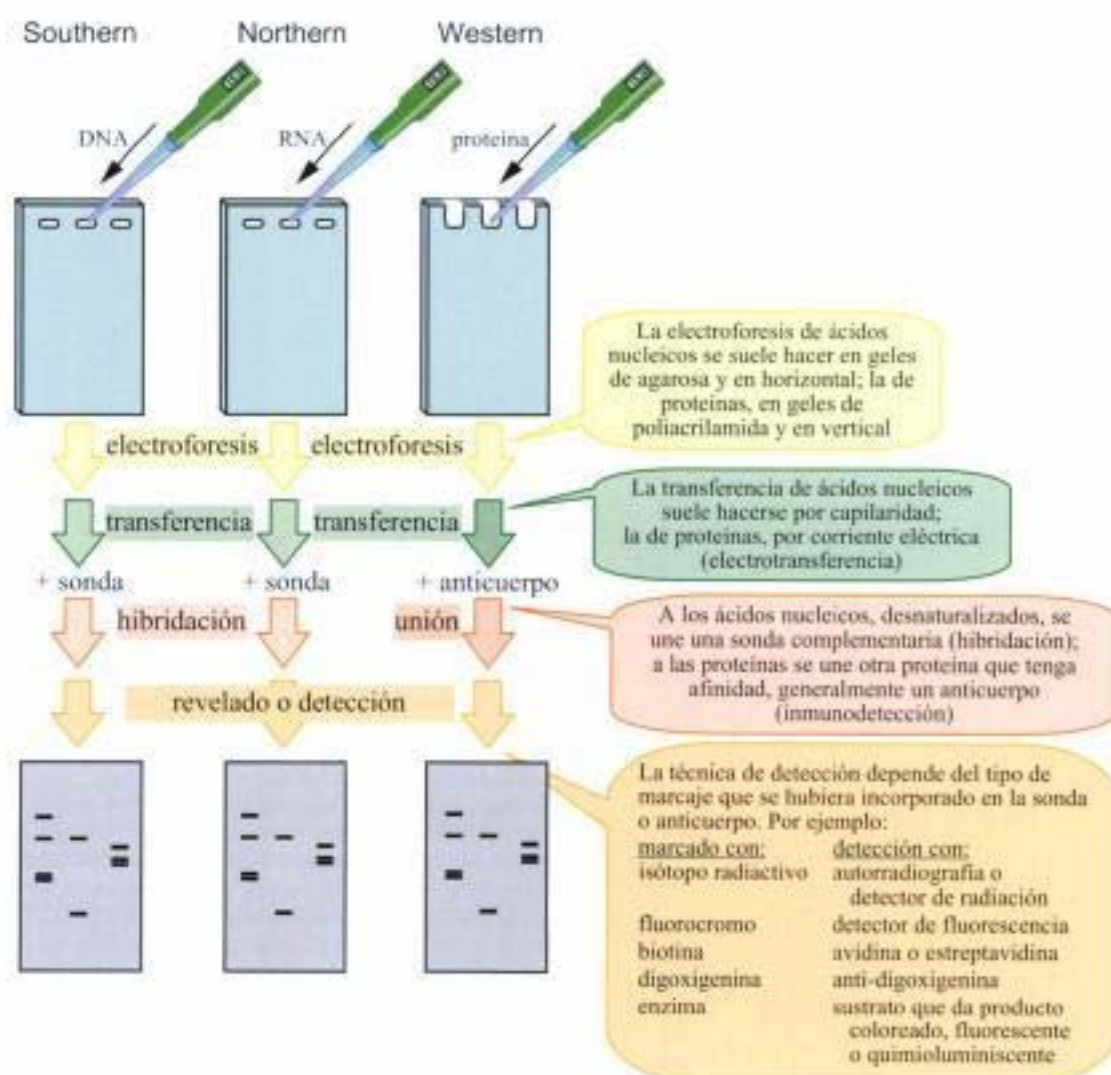
Uno de los métodos surgidos de la hibridación *in situ* cromosómica es el denominado **pintado de cromosomas** (*chromosome painting*), donde se usa como sonda una mezcla de fragmentos de DNA procedentes de un mismo cromosoma, con lo cual se observa fluorescencia en el cromosoma completo. Se facilita así la identificación visual de cada cromosoma (o varios simultáneamente, si se emplean sondas con distintos fluorocromos), en lo que se ha llamado un "cariotipo molecular". Esto es útil para estudiar grandes reorganizaciones de los cromosomas, que tienen lugar por ejemplo en procesos cancerosos.

13.3.4 Técnicas relacionadas

Se comenta finalmente una técnica de análisis que, aun no estando basada en los principios de hibridación, posee una similitud en los planteamientos que hace interesante su exposición en este lugar. Han surgido también otros numerosos métodos de análisis de ácidos nucleicos, variantes de los aquí expuestos, que combinan el reconocimiento de los ácidos nucleicos por sondas con la interacción específica entre aquéllos y proteínas, tales como los ensayos de retardo en gel o los métodos *Southwestern* y *Far-Western*. Su descripción excede el alcance de este libro.

13.3.4.1 Transferencia Western

A pesar de que en este caso no existe hibridación entre ácidos nucleicos, se describe aquí este método por su similitud técnica y mecanística con las hibridaciones Southern y Northern. El ensayo Western (western = del oeste) se emplea para detectar las **proteínas** presentes en una muestra. Para ello, son sometidas a electroforesis en gel de poliacrilamida (separándose en función de tamaño y carga, salvo en el caso frecuente de electroforesis en presencia de SDS, que separa sólo en función de tamaño) y transferidas a una membrana (western blot). La detección de las bandas proteicas en la membrana o filtro se realiza gracias a la unión específica de un anticuerpo contra la proteína o región proteica que se quiere estudiar. De forma similar al marcaje y detección de sondas de ácido nucleico, se puede marcar directamente el anticuerpo o detectarlo mediante la unión de una proteína ligante marcada a su vez, tanto empleando isótopos como técnicas no radiactivas. El paralelismo (y las diferencias) entre los 3 métodos se resume en la figura.



De forma análoga, también se realizan ensayos "dot-blot" y "slot-blot" para proteínas, con un planteamiento similar al expuesto para ácidos nucleicos, salvo por el tipo de muestra, y un procedimiento de unión y detección completamente análogo al del western (en este caso, obviamente, al no realizar electroforesis no es precisa la transferencia).

Tema 14

Preparación y marcaje de sondas para hibridación

14.1 Preparación de sondas genéticas	175
14.1.1 Clasificación de las sondas, en función del método de preparación	176
14.1.2 Sondas convencionales	176
14.1.2.1 Sondas de DNA obtenidas por clonación celular (tecnología del DNA recombinante)	176
14.1.2.2 Sondas de DNA obtenidas por clonación acelular (amplificación por PCR)	176
14.1.2.3 Sondas convencionales de RNA (ribosondas)	177
14.1.3 Sondas sintéticas	177
14.1.3.1 Síntesis química de oligonucleótidos	178
14.1.3.2 Diseño de una sonda sintética	179
14.2 Marcaje de las sondas	180
14.2.1 Marcaje directo en la molécula de sonda	181
14.2.1.1 Marcaje con DNA polimerasa	182
a) Método de traslado de la mella	182
b) Método de iniciación o cebado al azar	183
14.2.1.2 Marcaje terminal	183
a) Marcaje con polinucleótido quinasa	184
b) Marcaje terminal por relleno	184
14.2.2 Marcaje externo a la sonda	185
14.2.3 Marcaje indirecto	185
14.2.4 Tipos de grupos marcadores e indicadores y su detección	185
14.2.4.1 Marcadores radiactivos	185
14.2.4.2 Marcadores no radiactivos	186
14.2.4.3 Grupos indicadores	186

Como ya se ha indicado, la detección de ácidos nucleicos mediante ensayos de hibridación depende de la presencia de la secuencia diana a detectar en la muestra y de la disponibilidad de una **sonda** u oligonucleótido de secuencia complementaria a la diana. Por tanto, la utilidad de la hibridación, el punto clave, radica en la sonda: de su naturaleza depende que el ensayo sea fructífero. El papel de la sonda equivale al del anticuerpo en un inmunoensayo y, al estar marcada, es posible localizar y cuantificar cualquier secuencia diana con la cual se pueda hibridar. Gracias al empleo de sondas, la sensibilidad de la hibridación es muy elevada, hasta un millón de veces superior a la detección directa por otros métodos, por ejemplo de tinción tras electroforesis. La sonda se convierte así en un elemento esencial para cualquier estudio molecular, en especial para analizar la variabilidad o polimorfismo del genoma. De ahí la conveniencia de estudiar los métodos de preparación y marcaje de sondas, como aspectos esenciales para su empleo en ensayos de hibridación.

14.1 PREPARACIÓN DE SONDAS GENÉTICAS

Se utiliza una gran variedad de métodos para obtener sondas, tanto de DNA (en hebra doble y hebra sencilla) como de RNA. Obviamente, para su empleo en el ensayo de hibridación la sonda siempre ha de encontrarse en forma de hebra sencilla, para lo cual se acude a la simple desnaturalización de las sondas de doble hebra.

14.1.1 Clasificación de las sondas, en función del método de preparación

Las sondas se pueden clasificar según varios criterios: longitud o tamaño de la secuencia diana, método de marcaje y detección, tamaño de la sonda, método de preparación, etc. Este último criterio es el más utilizado, distinguiéndose entre sondas preparadas por métodos convencionales (por clonación celular o acelular, temas 15 y 16) y por métodos de síntesis química.

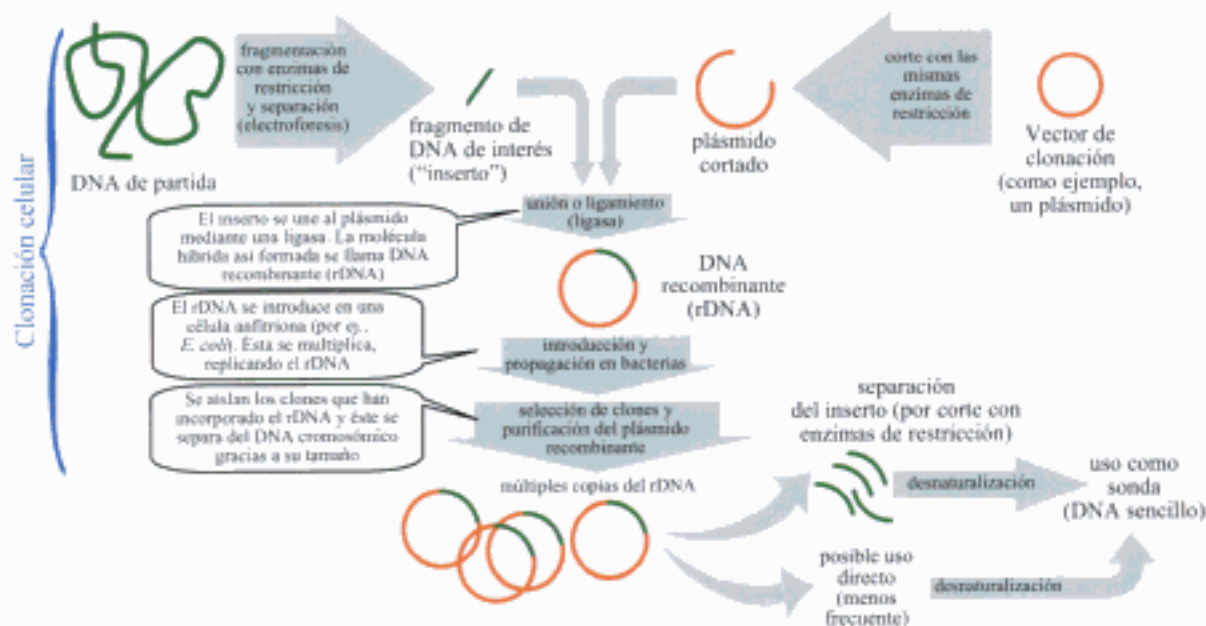
Tipo de sonda:	Convencionales			Sintéticas
Material de partida:	DNA de doble hebra	RNA	DNA de doble hebra	Nucleótidos
Método de obtención:	Clonación celular clásica	Clonación acelular (PCR)	PCR con transcriptasa inversa (RT-PCR)	Clonación celular en vector especial seguida de transcripción
Naturaleza de la sonda formada:	Fragmento de DNA (oligodesoxirribonucleótido)		"Ribosonda": fragmento de RNA (oligorribonucleótido)	Oligonucleótido ("oligo")

14.1.2 Sondas convencionales

También llamadas **sondas genómicas, clonadas o biológicas**, fueron las primeras utilizadas y han contribuido especialmente al progreso de la Biología Molecular.

14.1.2.1 Sondas de DNA obtenidas por clonación celular (tecnología del DNA recombinante)

Son de tamaño (longitud) variable, entre 100 bases y cientos de kb. Aunque los detalles experimentales de la clonación celular se estudian más adelante (tema 16), se adelantan aquí brevemente las etapas principales del proceso.



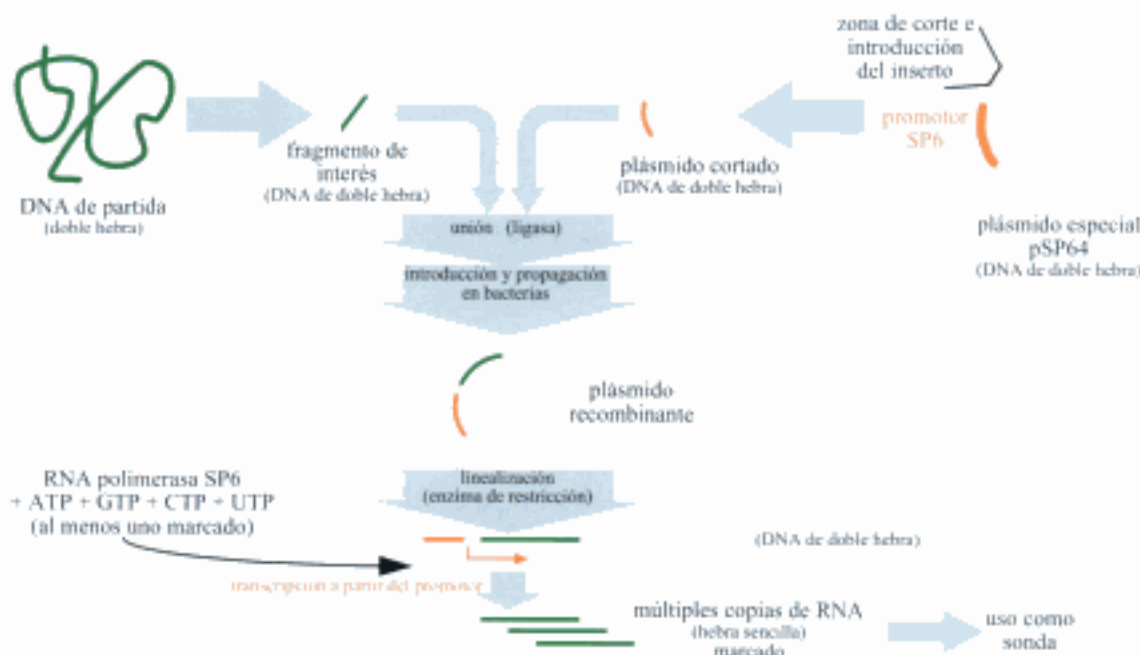
14.1.2.2 Sondas de DNA obtenidas por clonación acelular (amplificación por PCR)

Estas sondas son de menor tamaño que las obtenidas por clonación celular, entre 0,1 y 20 kb. El método de preparación (PCR, tema 15) también consta de un menor número de fases, por lo que las sondas se obtienen de forma más rápida y sencilla. En general, partiendo de DNA de doble hebra se preparan sondas de la misma naturaleza, salvo en la PCR asimétrica, que conduce a DNA de hebra sencilla, y en la RT-PCR, que permite obtener sondas de DNA de doble hebra partiendo de un RNA (pág. 195).



14.1.2.3 Sondas convencionales de RNA (ribosondas)

La clonación celular se puede emplear también para preparar sondas de RNA de unos miles de nucleótidos. El proceso, relativamente complejo, se basa en la transcripción de un DNA previamente clonado. Se usa un plásmido especial (generalmente el pSP64), que se caracteriza por incluir el promotor de una RNA polimerasa de fago (*promotor SP6*). El rDNA resultante, una vez purificado y linealizado, se utiliza como molde para la transcripción por RNA polimerasa SP6, que transcribe el inserto específicamente a partir de la secuencia promotora SP6, sintetizando numerosas copias (de hebra sencilla) de RNA. Si alguno de los NTPs, sustratos de la RNA polimerasa, está marcado (pág. 181), el RNA se puede utilizar directamente como sonda; si no, deberá marcarse el RNA posteriormente a su obtención.



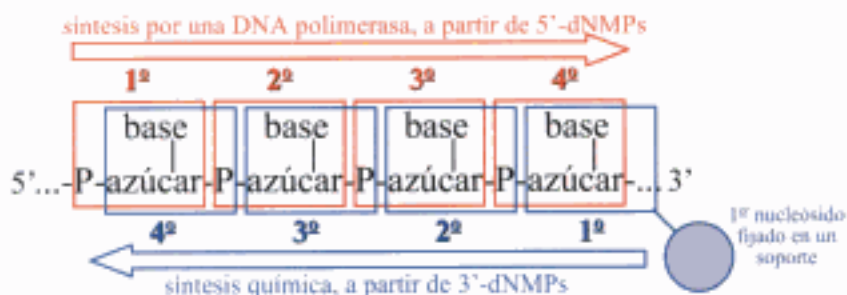
Ventajas de las ribosondas	Desventajas de las ribosondas
Al no poderse formar el homodúplex sonda:sonda, se favorece la eficacia de la interacción de la sonda con la secuencia diana.	Es necesario usar un vector de clonación particular, con un promotor de transcripción.
Las sondas no unidas específicamente pueden ser eliminadas por digestión con RNasa, que degrada hebras sencillas, pero respeta los híbridos RNA sonda : DNA diana.	Se aumenta la complejidad al precisarse la etapa de reacción RNA polimerasa.
Se puede emplear como control negativo la sonda complementaria (secuencia en dirección <i>antisentido</i> , pág. 248, conseguida al incorporar el inserto junto al promotor en orientación contraria).	Puesto que el RNA es mucho más lábil que el DNA, y la RNasa es una contaminación muy frecuente, se requiere un manejo cuidadoso de las sondas de RNA para evitar su degradación y el uso de técnicas y material estériles y específicamente libres de RNasa.

14.1.3 Sondas sintéticas

La capacidad de síntesis química es esencial en Ingeniería Genética no sólo para preparar sondas, sino también para modificar parcial o grandemente un producto génico o para sintetizar genes completos. En general, se obtienen sondas de DNA más cortas que por clonación. El método más popular es la síntesis de sondas cortas de DNA (de 15 a 45 nucleótidos), monocatenarias, llamadas **oligonucleótidos** o simplemente **oligos**. Éstos pueden ser también de RNA, aunque se sintetizan con mayor dificultad.

14.1.3.1 Síntesis química de oligonucleótidos

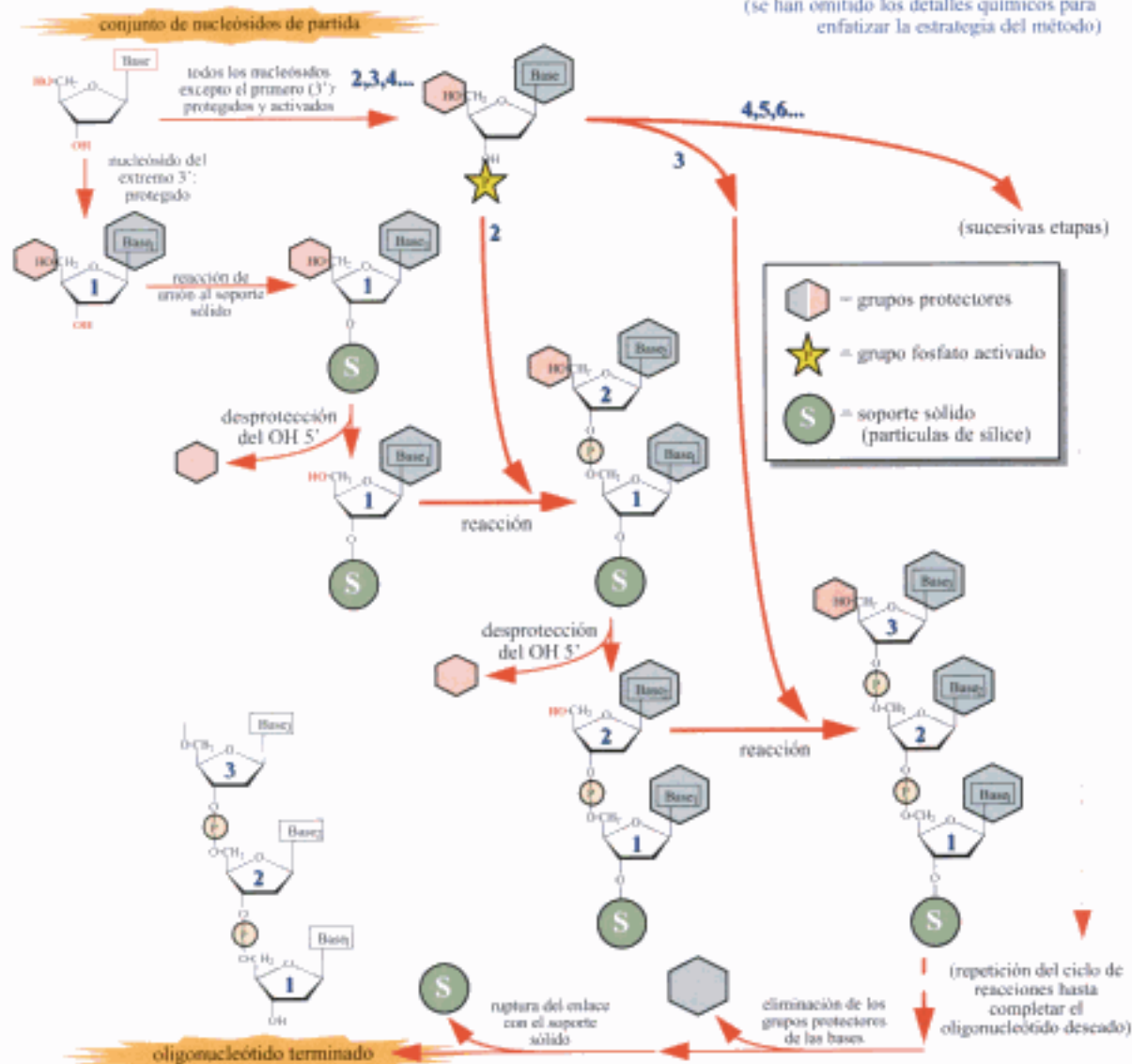
El aspecto básico de la síntesis es la formación repetitiva de un enlace éster entre el grupo 3'-fosfato de un nucleótido y el grupo 5'-OH de otro, para dar un enlace fosfodiéster. Es decir, se parte de 3'-NMPs y se avanza en sentido 3'→5', a diferencia de la síntesis de ácidos nucleicos en las células (tanto replicación como transcripción emplean 5'-NTPs y avanzan de 5' a 3'). Los NMPs se añaden uno a uno partiendo del primer nucleósido, que está unido por su posición 3' a un soporte sólido y constituirá el extremo 3' del ácido nucleico una vez sintetizado.



El principal problema es la gran reactividad de los dNMPs, con grupos -OH primario y secundario, un -NH₂, un grupo fosfato y grupos amino y/o carbonilo en la base nitrogenada, por lo que se requieren métodos de bloqueo y desbloqueo que no afecten irreversiblemente a ninguno de los componentes de la molécula de nucleótido. Por otro lado, el grupo fosfato (en 3') ha de activarse para que reaccione con el -OH (en 5' del nucleótido anterior).

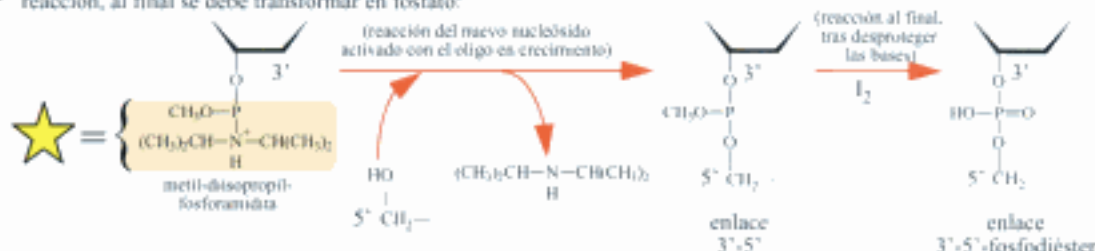
Método desarrollado en los años 70 por H. Gobind Khorana para la síntesis química de oligonucleótidos:

(se han omitido los detalles químicos para enfatizar la estrategia del método)

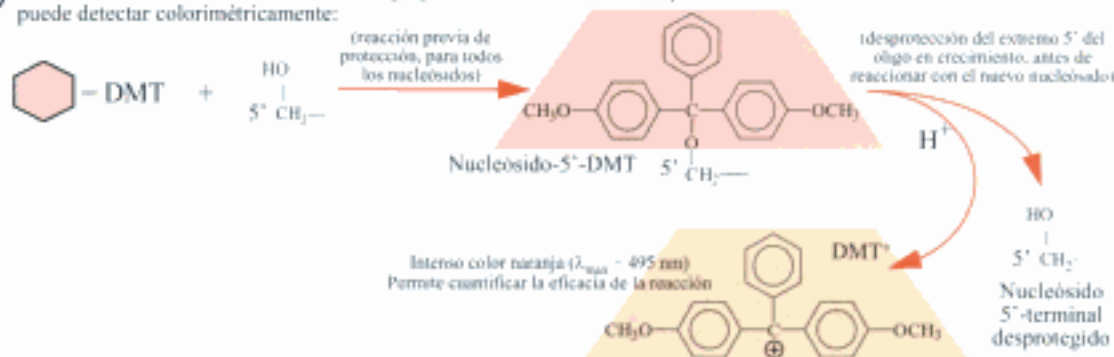


Detalles (reacciones, grupos protectores y activantes) de la síntesis de oligos por el método de Khorana:

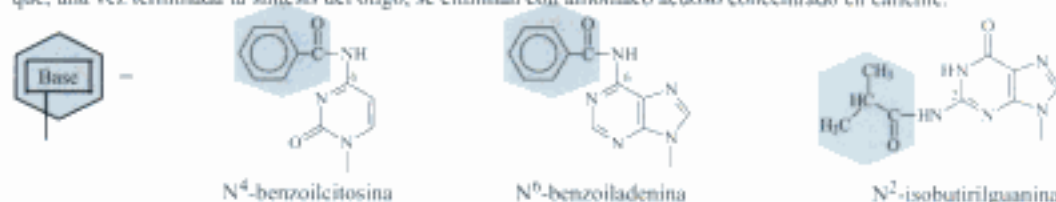
- 1 Activación del fosfato 3': no se introduce como tal, sino en forma de otro grupo que, además de activar la reacción, al final se debe transformar en fosfato:



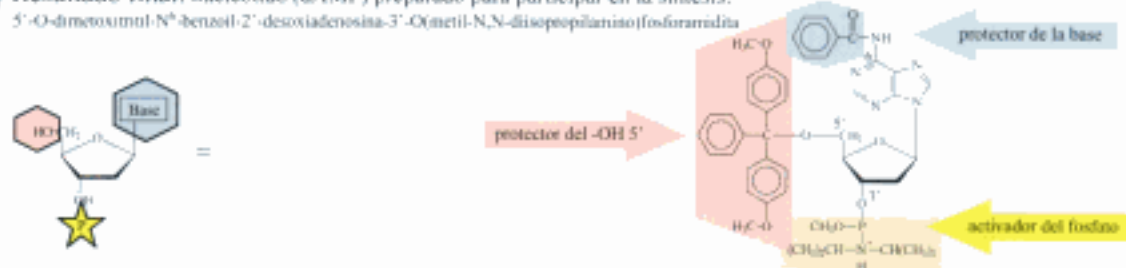
- 2 Protección del -OH 5': mediante un grupo dimetoxitritilo (DMT), que se libera en medio ácido. Esta liberación se puede detectar coloriméticamente:



- 3 Protección de los grupos amino exocíclicos de las bases: habitualmente con grupos benzoilo o isobutirilo que, una vez terminada la síntesis del oligo, se eliminan con amoníaco acuoso concentrado en caliente:



- 4 Resultado final: nucleótido (dAMP) preparado para participar en la síntesis:



Estos procesos están ya hoy día automatizados, son eficientes y de bajo costo. Debido a la que la longitud de estos oligos se ve limitada por el rendimiento de las reacciones de síntesis, la sensibilidad alcanzada con ellos es, en general, menor que con las sondas obtenidas por clonación. Sin embargo, la continua mejora en los métodos de síntesis química permite la obtención de oligos cada vez mayores.

Respecto a la química de síntesis de oligonucleótidos de RNA, está más atrasada debido a la dificultad de proteger el grupo OH 2' sin afectar al 3'.

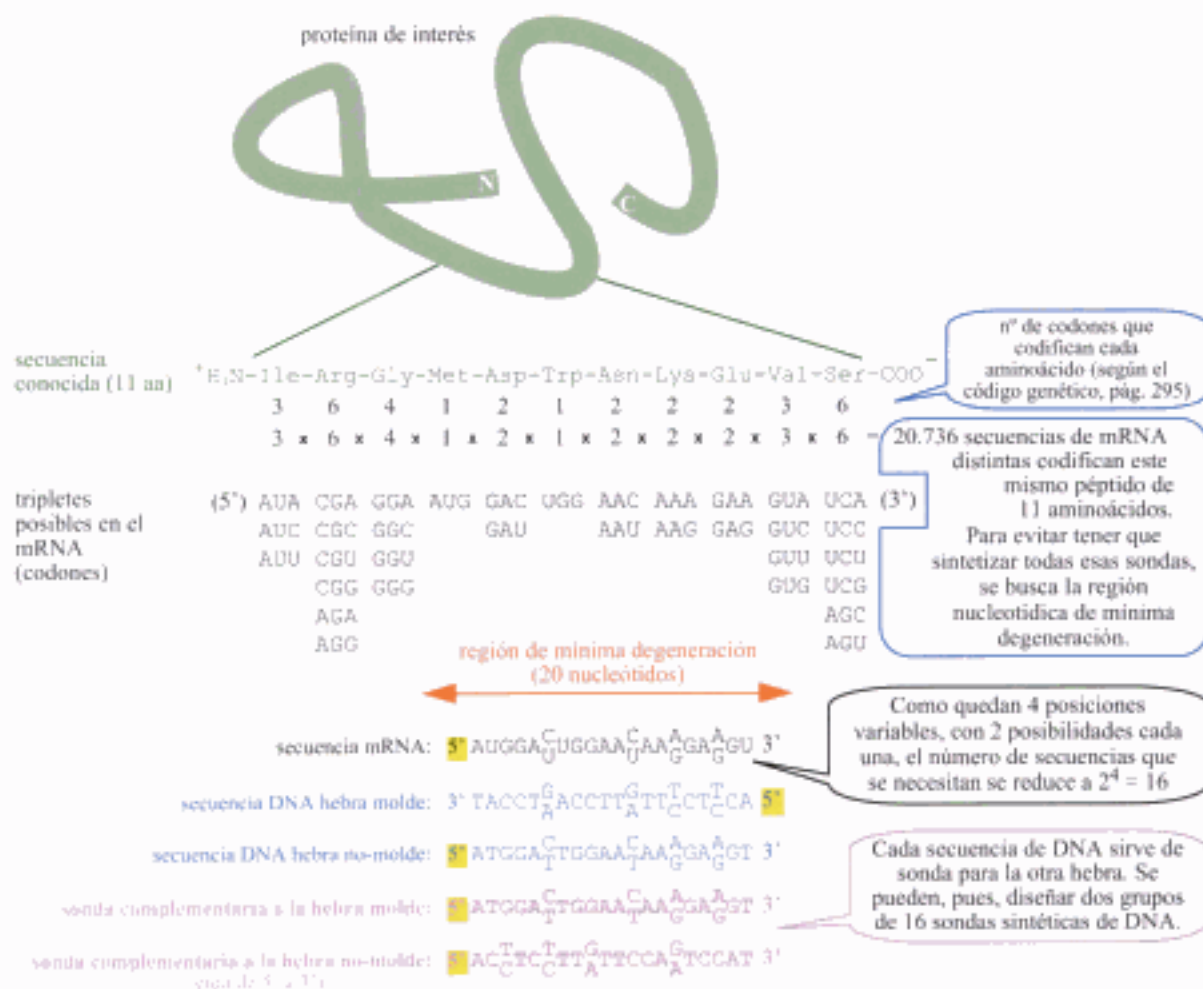
14.1.3.2 Diseño de una sonda sintética

Las sondas sintéticas se diseñan de forma que contengan la secuencia complementaria a una zona del DNA diana, por ejemplo, de una parte del gen que interesa analizar. El objetivo es preparar un oligonucleótido que pueda hibridarse con dicho DNA. Generalmente, se pretende localizar el gen que codifica la síntesis de una proteína determinada. Para diseñar la sonda, es necesario:

1. Conocer la secuencia completa de la proteína, o al menos de un fragmento de ella, al que corresponderá la secuencia diana de DNA.
2. A partir de esa secuencia aminoácida, se establece (en base al código genético) la secuencia nucleotídica que habrá de tener el DNA. Sin embargo, debido a la degeneración del código genético (distintos trinucleótidos codifican un mismo aminoácido, pág. 295), hay un número elevado de secuencias de DNA que pueden codificar el

mismo polipéptido y, obviamente, no se puede conocer *a priori* cuál es la que lo hace en el genoma real de la célula, por lo que es preciso sintetizar sondas que cubran todas las posibilidades.

3. Para reducir su número, se selecciona el segmento de secuencia que posea la menor degeneración posible.
4. Una vez sintetizadas todas las sondas, se ensaya su hibridación con el DNA de partida; una de ellas será la correcta.



En este tipo de experimentos, los métodos de síntesis y secuenciación de proteínas y de DNA convergen y se complementan. Como se acaba de ver, la secuencia de una proteína aislada puede usarse para aislar el gen correspondiente; de una forma similar, aunque inversa, un gen clonado (y, por tanto, secuenciado) puede usarse para aislar la proteína producto: la secuencia del gen permite diseñar y sintetizar químicamente un péptido corto que representa una parte del producto proteico del gen. Entonces, una de las posibilidades es, por ejemplo, utilizar anticuerpos contra este péptido sintético para identificar y purificar la proteína para su estudio posterior.

14.2 MARCAJE DE LAS SONDAS

El empleo de sondas de DNA o RNA como herramientas analíticas para localizar secuencias diana por ensayos de hibridación requiere obligatoriamente un proceso de marcaje de la sonda, que permita su detección. La molécula o grupo **marcador** puede tener diversa naturaleza, siempre y cuando se pueda detectar su presencia de forma **directa** o **indirecta**; los marcadores más utilizados son **isótopos radiactivos**, **fluorocromos** y **enzimas**. En general, dan lugar a métodos sencillos de realizar, que detectan concentraciones muy bajas de la sonda. El marcaje debe ser estable bajo diversas condiciones de temperatura, detergentes, disolventes, etc., y no afectar a la reacción de hibridación. Los métodos radiactivos son los empleados clásicamente, mientras que los otros marcadores surgen principalmente para evitar los riesgos derivados de la radiación, para aprovechar la mayor vida media de sus reactivos de marcaje y, en algunos casos, para mejorar el límite de detección.

De forma genérica, pueden plantearse dos tipos de proceso de marcaje:

- **In vivo** (más correctamente, en cultivo): Se hace suministrando nucleótidos radiactivos a células en cultivo, que los incorporan en sus ácidos nucleicos de nueva síntesis. Su uso es muy limitado, por ejemplo, para preparar DNA viral marcado en células infectadas por virus.

- *In vitro*: Es más versátil. Puede hacerse empleando isótopos radiactivos o por otros métodos alternativos. Se describen a continuación algunas de las diversas variantes metodológicas.

En muchas ocasiones, en especial con ribosondas y sondas sintéticas, la sonda se marca **al mismo tiempo que se prepara** incluyendo, con los NTPs que hay que suministrar a la reacción de síntesis, alguno de ellos marcado (con un isótopo radiactivo, un fluorocromo o cualquier tipo de grupo indicador) mientras que en otras el marcaje se realiza **tras haber completado la síntesis** de la sonda. Puesto que la preparación de sondas ya se ha estudiado, nos centraremos ahora en este último caso, aunque las consideraciones relativas a la detección sirven igualmente para sondas marcadas durante su obtención.

El principio básico del marcaje es la unión química o *conjugación* del **marcador** a la sonda (en el caso radiactivo, en realidad el isótopo forma parte de la molécula de sonda, no se une a ella). Como se acaba de indicar, esto puede hacerse uniendo el marcador a la sonda ya preparada o a un nucleótido que se incorporará durante la preparación de aquella. Por otra parte, la sonda puede llevar unido el marcador detectable (*métodos directos* de marcaje, a continuación) o bien otra molécula no detectable o "**grupo indicador**" (en inglés, *reporter*), con la capacidad de ser reconocido posteriormente por una "**proteína detectora**", que es la que va marcada (*métodos indirectos*, pág. 185).

Marcaje directo: el marcador, que se puede detectar, va unido a la sonda.

Dos alternativas, después de la síntesis de la sonda o durante ella:



Marcaje indirecto: el "marcador" unido a la sonda no es detectable directamente; se le llama indicador.

Para la detección, el indicador se une a una proteína detectora, que lleva el marcaje.

Dos alternativas, después de la síntesis de la sonda o durante ella:



14.2.1 Marcaje directo en la molécula de sonda

El grupo marcador se puede unir covalentemente a grupos funcionales específicos o formar enlaces covalentes de forma inespecífica. Los ácidos nucleicos muestran menos grupos funcionales adecuados para este fin que las proteínas; básicamente, algunos átomos de las bases nitrogenadas (especialmente, aquéllos no implicados en el apareamiento con la base complementaria) y los grupos terminales 3'-OH y 5'-P.

Marcaje covalente de la sonda (por métodos químicos):



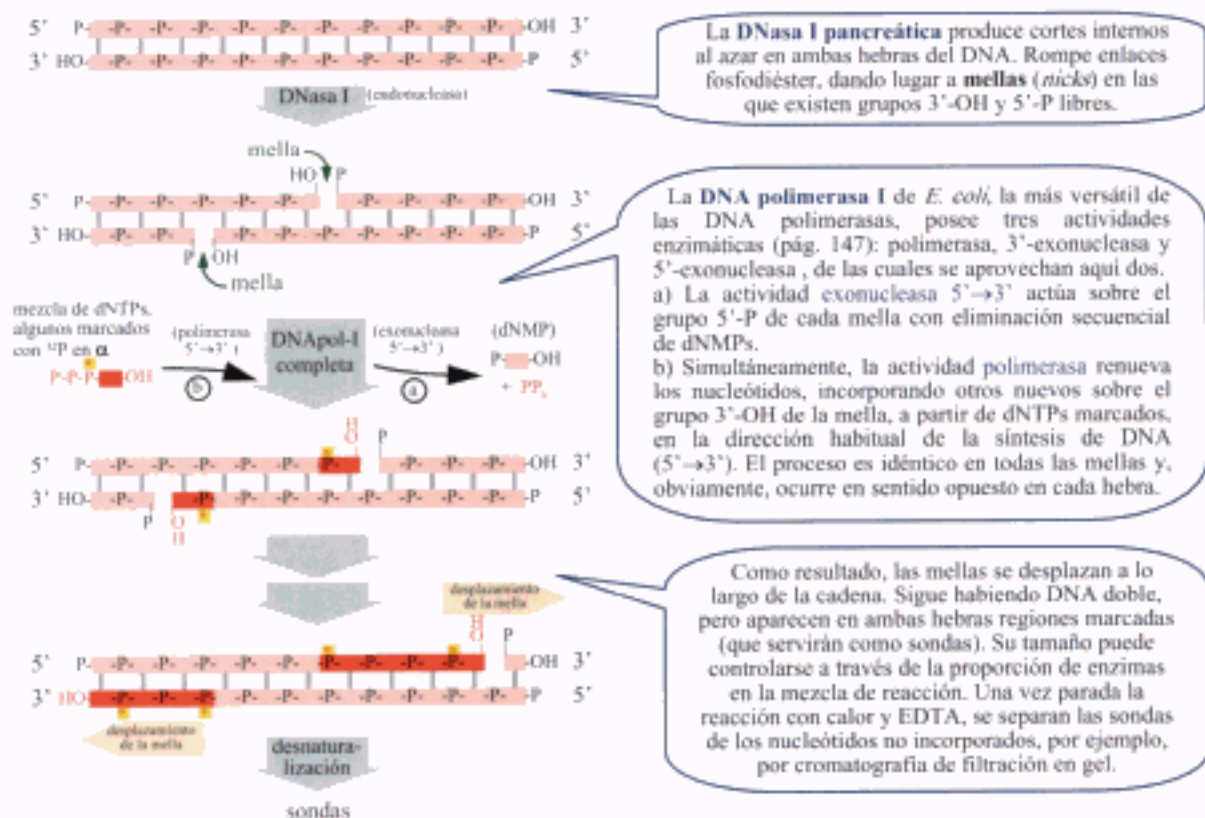
En esta categoría de métodos pueden considerarse tanto la inclusión de nucleótidos marcados durante la síntesis de la sonda como el marcaje de la sonda ya terminada. De estos últimos métodos, se comentan a continuación los más representativos.

14.2.1.1 Marcaje con DNA polimerasa

Se trata de métodos de **síntesis enzimática**, los más comunes. Los cuatro tipos de nucleótido (al menos uno de ellos marcado) se incorporan durante la reacción de síntesis de una cadena de DNA. Existen dos variantes: traslado de la mella (*nick translation*) e iniciación o cebado al azar (*random priming*).

a) Método de traslado de la mella

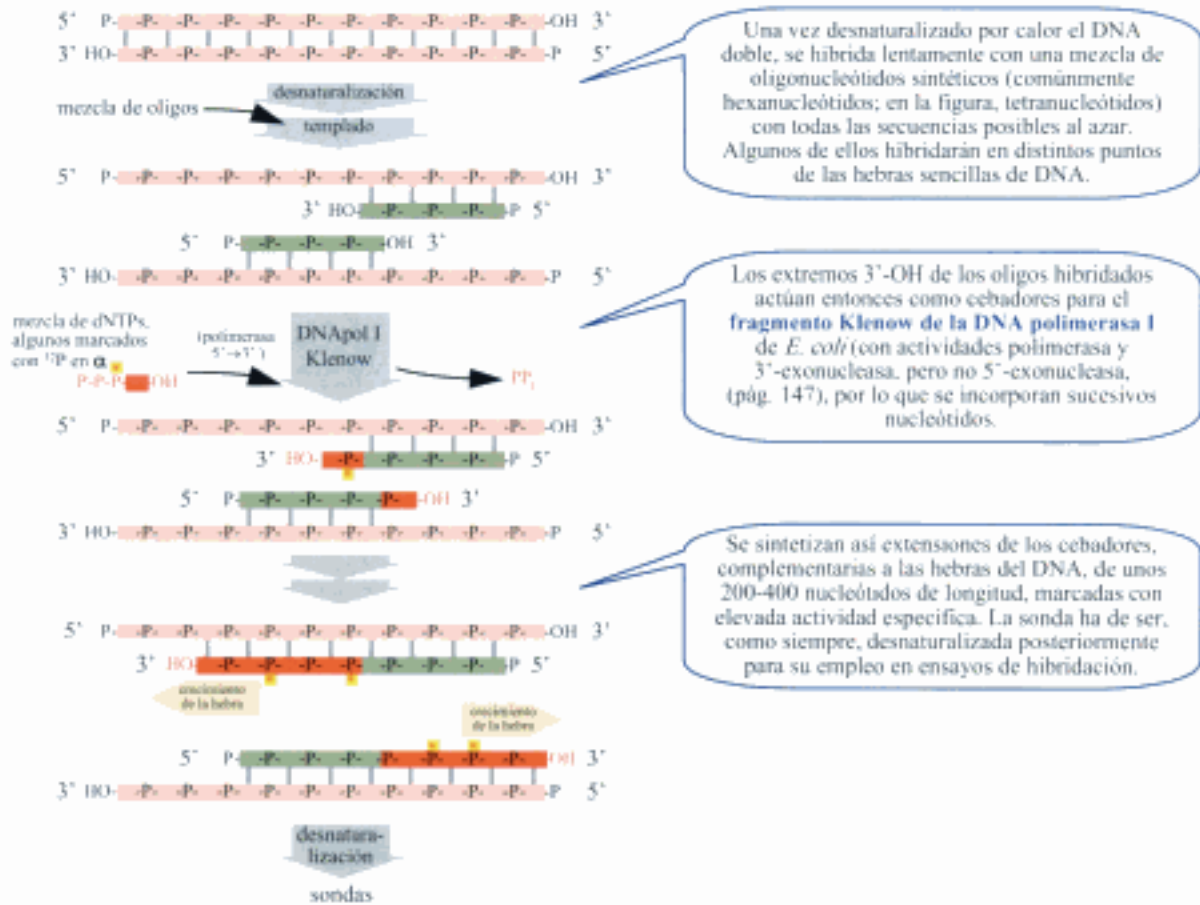
Es un método fácil y rápido, considerado el ejemplo clásico de incorporación enzimática de un marcaje, sin necesidad de desnaturalización de la sonda de DNA doble. Requiere dos enzimas distintas: una desoxirribonucleasa (DNasa I pancreática) y una DNA polimerasa (DNApol-I de *E. coli*). El resultado del marcaje es la sustitución de una gran parte (30-60%) de los nucleótidos originales de cada hebra de DNA por otros marcados radiativamente en su posición α -5'-P.



Puesto que el nucleótido marcado se incorpora a la sonda, en lugar del marcaje radiactivo (en la figura) pueden emplearse nucleótidos marcados con fluorocromos o con grupos indicadores (págs. 181, 185).

b) Método de iniciación o cebado al azar

Esta alternativa, a diferencia de la anterior, requiere la desnaturalización inicial del DNA doble, para que hibride un oligonucleótido que actúa como cebador de la replicación por DNAPol-I de *E. coli*. Ahora bien, en este caso sólo se requiere la actividad polimerasa, por lo que se prescinde del *fragmento pequeño* y se suele utilizar el *fragmento de Klenow* (pág. 147). Esta técnica ofrece una serie de ventajas de carácter técnico respecto a la anterior, cuya consideración escapa a nuestros objetivos de carácter esencialmente conceptual.



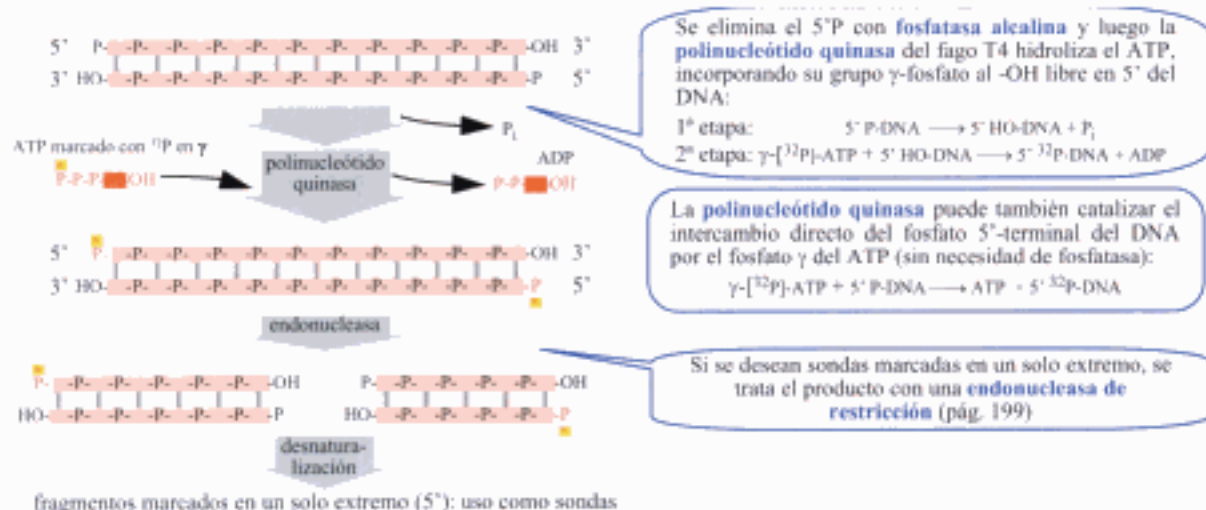
Al igual que el método anterior, admite también nucleótidos marcados con fluorocromos o con grupos indicadores.

14.2.1.2 Marcaje terminal

En este caso se aplica, en lugar de una síntesis, una **modificación enzimática** de las hebras de ácido nucleico existentes, mediante la incorporación de uno o unos pocos nucleótidos marcados a uno o ambos extremos de la sonda de DNA o RNA. Aunque se utilizan menos que los anteriores, son métodos útiles en ciertos casos, por ejemplo para marcar oligos de hebra sencilla. Se utilizan dos planteamientos; el marcaje con polinucleótido quinasa y el marcaje terminal por relleno (*fill-in end-labeling*).

a) Marcaje con polinucleótido quinasa

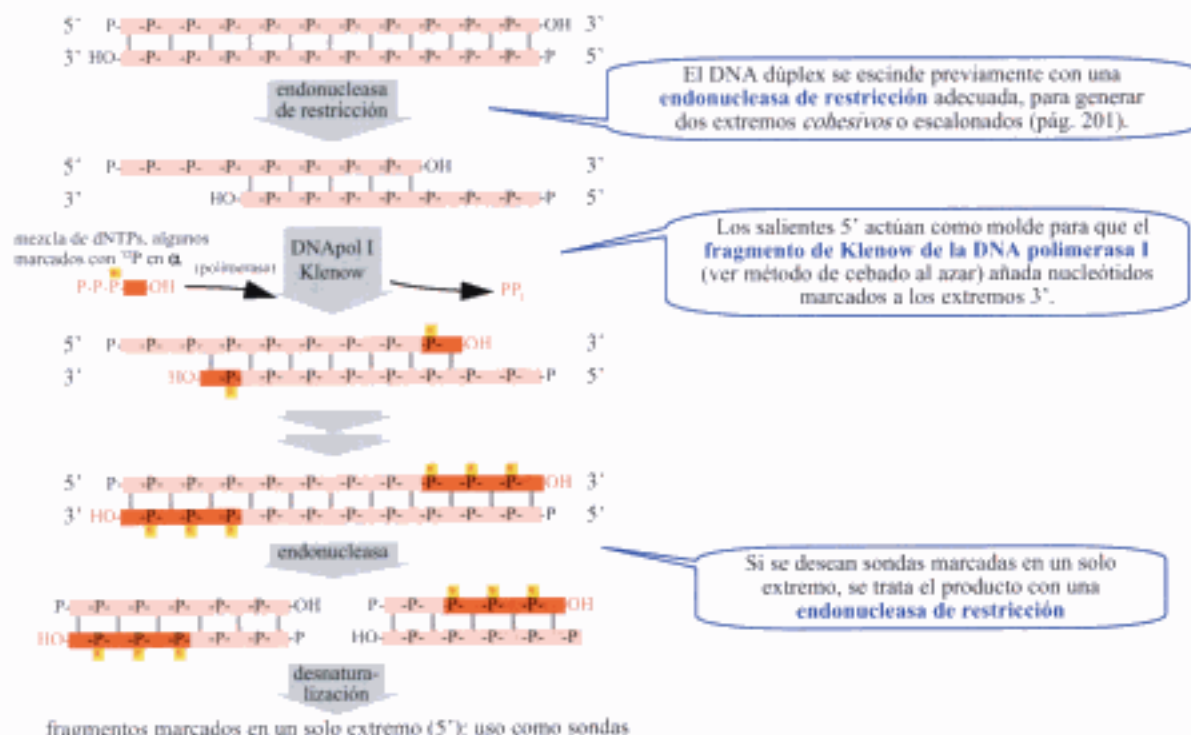
Este método no depende de la existencia de una hebra molde y, por tanto, puede emplearse para cualquier tipo de sonda de DNA, sencillo o doble, o de RNA. Consiste, en esencia, en un intercambio del 5'-P terminal de cualquier hebra por el fosfato (marcado) en posición γ de un dNTP, con la intervención de polinucleótido quinasa y, opcionalmente, fosfatasa alcalina.



En este caso sólo pueden utilizarse marcajes radiactivos, pues lo único que se incorpora a la sonda es el fosfato γ del ATP.

b) Marcaje terminal por relleno

Este método se caracteriza por ser el único que utiliza una enzima de restricción específica para formar dos extremos cohesivos, que luego son "rellenados" por la DNApol-I.



Este método de nuevo permite el uso tanto de nucleótidos radiactivos (en la figura) como de aquéllos marcados con fluorocromos o con grupos indicadores.

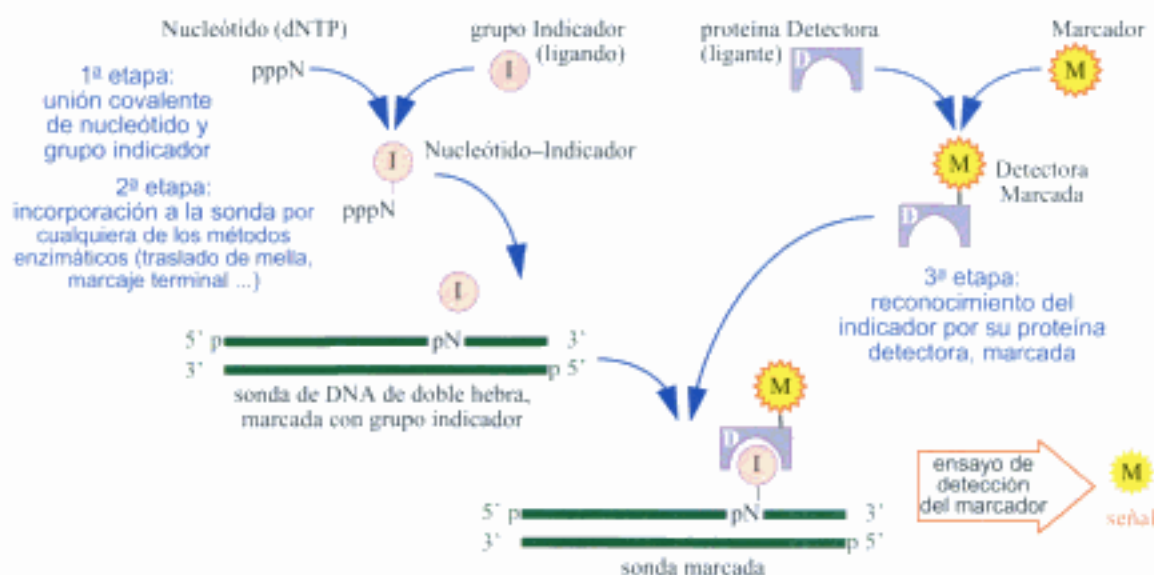
14.2.2 Marcaje externo a la sonda

En lugar de marcar sus propios nucleótidos, la sonda se puede unir con otra molécula que a su vez lleva el marcador. Una posibilidad es la unión de proteínas mediante reactivos entrecruzantes (crosslinkers: formaldehído, glutaraldehído, etc.). También puede unirse el marcador por interacción electrostática con el DNA o emplearse agentes intercalantes, que se insertan entre los pares de bases de la doble hélice (híbrido sonda : DNA) de forma no covalente.



14.2.3 Marcaje indirecto

Como se ha indicado (pág. 180), puede convenir que a la sonda se una un grupo no detectable o indicador, el cual en una segunda etapa es reconocido por otra molécula, generalmente una proteína, que incorpora el marcador detectable. Para la unión del indicador son aplicables las consideraciones ya realizadas respecto a la unión del marcador en los métodos directos. Se ilustra un ejemplo correspondiente a la modificación de los NTPs:



14.2.4 Tipos de grupos marcadores e indicadores y su detección

14.2.4.1 Marcadores radiactivos

El marcaje radiactivo o isotópico es el que se ha empleado con más frecuencia, aunque va siendo desplazado progresivamente por los métodos no radiactivos. Aunque no existe consenso acerca del marcaje ideal, en la mayoría de los casos se hace con ^{32}P . Puesto que la corta vida media de este isótopo (14,2 días) crea dificultades en la preparación, comercialización y uso rutinario de los ensayos, se utilizan también otros alternativos (^{35}S , ^3H , ^{125}I o ^{14}C).

La detección y registro de la hibridación se hace por *autorradiografía*: el soporte con el material hibridado se pone en contacto con una película fotográfica (del tipo empleado para las radiografías), la cual queda impresionada con una imagen que muestra la presencia de la radiactividad (ver, por ejemplo, el método de Southern, pág. 171). La intensidad de la señal depende de la intensidad de la radiación emitida por el isótopo y del tiempo de exposición.

14.2.4.2 Marcadores no radiactivos

En función del marcador empleado, la detección se hace por fluorimetría o añadiendo un sustrato específico cuya transformación enzimática por la enzima marcadora pueda detectarse por alguno de los siguientes principios:

- Espectrofotometría: medida de la absorción de la luz por los productos de la reacción, bien solubles o insolubles, en el ultravioleta o el visible (colorimetría).
- Fluorimetría: medida de la luz emitida por productos de la reacción que posean un fluoróforo (pág. 126).
- Quimioluminiscencia: medida de luz emitida como consecuencia de reacciones con esta característica (por ejemplo, la transformación de las luciferinas, catalizada por luciferasa, o la oxidación del luminol).

Debe señalarse que estos y otros procedimientos no son exclusivos del marcaje de sondas, sino que se aplican igualmente en la detección de proteínas, antígenos, pequeños ligandos, etc.

Resumen de marcadores, indicadores y métodos para su detección

 Grupo indicador (ligando)	 Proteína detectora o ligante (*)	 Marcador	Detección del marcador	Sustrato de la reacción enzimática de detección
Biotina	Avidina, Estreptavidina, o variantes (**)	Isótopo radiactivo	Autoradiografía	
Digoxigenina	Anticuerpo anti-DIG	Fluorocromo	Fluorimetría	
		Peroxidasa (abreviada como PO, POD, HRPO, HRP)	Reacción Cromogénica:	ABTS TMB Derivados de α -naftol
			Reacción Fluorogénica:	Derivados de dioxetano o de luminol
		Fosfatasa Alcalina (AP)	Reacción Cromogénica:	NBT+BCIP NPP
			Reacción Fluorogénica:	AttoPhos™ HNPP+FastRedTR
			Quimioluminiscencia:	Derivados de dioxetano
		β -Galactosidasa (β -Gal)	Reacción Cromogénica:	X-gal CPRG
		Oro coloidal	Microscopia	

* La proteína ligante, con afinidad por el grupo indicador, es marcada para su detección, por alguno de los métodos indicados en la tabla de la derecha.

** Variantes de avidina obtenidas por ingeniería genética para modificar la afinidad, reducir la señal inespecífica, etc.

Abreviaturas:

ABTS = 2,2'-azino-bis-(3-etilbenzotiazolinasulfonato) (6)

TMB = 3,3',5,5'-tetrametilbenzidina

NBT = azul de nitrotetrazolio

BCIP = 5-bromo-4-cloro-3-indolilfosfato

NPP = 4-nitrofenilfosfato

HNPP = 2-hidroxi-3-naftoico-2'-fenilamida fosfato

Fast Red TR = 4-cloro-2-metilbenzeno-diazotio

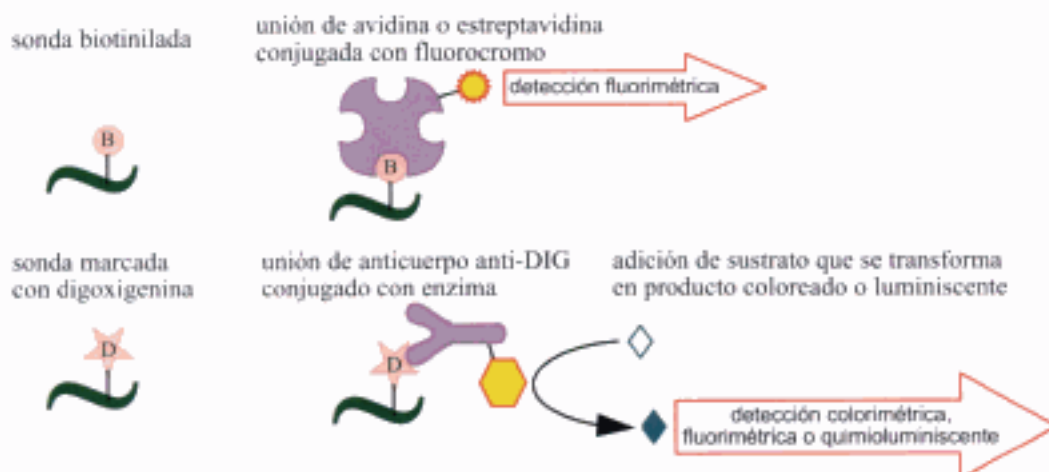
X-gal = 5-bromo-4-cloro-3-indolil- β -D-galactopiranosido

CPRG = rojo de clorofenol- β -D-galactopiranosido

14.2.4.3 Grupos indicadores

Los grupos más comúnmente usados son la **biotina** (vitamina natural) y la **digoxigenina** (esteroide de la planta *Digitalis*). Ambos son fácilmente conjugables a los nucleótidos y se unen con gran afinidad, respectivamente, a las proteínas avidina o estreptavidina y a anticuerpos anti-digoxigenina. Al igual que para los derivados fluorescentes, se dispone comercialmente de dNTPs ya unidos tanto a biotina como a digoxigenina.

La detección de las sondas que llevan biotina o digoxigenina se realiza mediante ensayos de unión, similares a los empleados para el inmunoensayo. A este fin, la proteína detectora, con afinidad por el grupo indicador, se marca previamente con un isótopo radiactivo, un grupo fluorescente o una enzima.



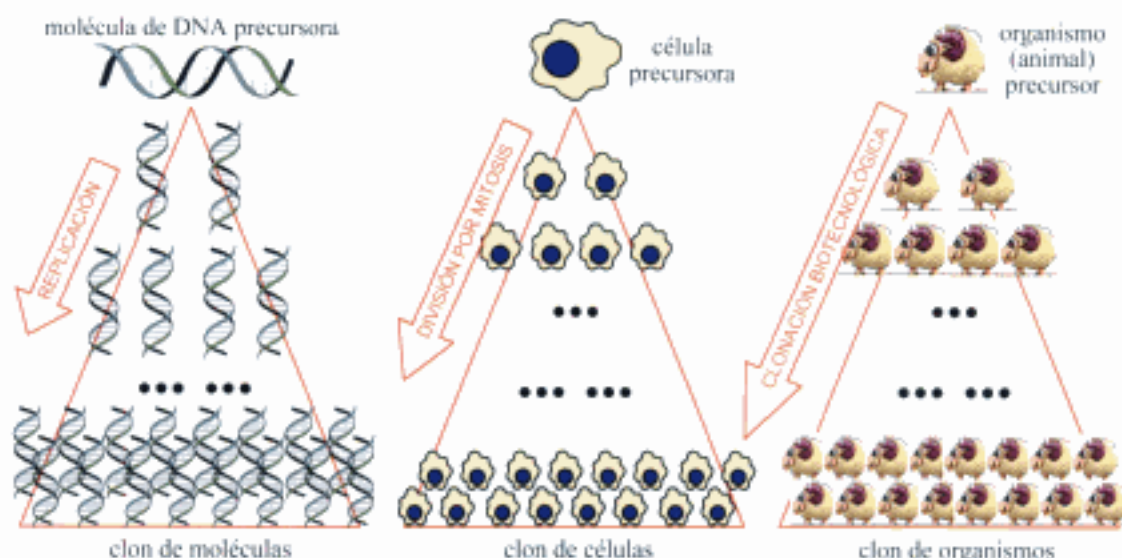
Tema 15

Clonación acelular: Reacción en cadena de la polimerasa (PCR)

15.1	Introducción general a la clonación.....	187
15.2	Amplificación <i>in vitro</i> del DNA: Reacción en cadena de la polimerasa o PCR	188
15.2.1	Principio del método	188
15.2.2	Etapas del proceso.....	189
15.2.3	Algunas características del proceso.....	192
15.2.4	Automatización de la PCR.....	193
15.2.5	Variantes del método.....	193
15.2.5.1	PCR "larga".....	193
15.2.5.2	PCR "anidada".....	193
15.2.5.3	PCR inversa	194
15.2.5.4	PCR con adaptadores	195
15.2.5.5	PCR asimétrica.....	195
15.2.5.6	RT-PCR: amplificación de RNA.....	195

15.1 INTRODUCCIÓN GENERAL A LA CLONACIÓN

Conceptualmente, el proceso de **clonación** (**clonado** o **clonaje**) consiste en la obtención de un **clon**, entendido como un conjunto de "elementos" genéticamente idénticos entre sí y a su precursor. Estos "elementos" pueden ser moléculas, células, tejidos, órganos u organismos pluricelulares completos. Cada componente individual de un clon contiene la misma información genética, el mismo genotipo, que el elemento de partida; por ello, se puede considerar



En cada caso, clon = conjunto de elementos genéticamente idénticos (entre sí y al elemento precursor)

que la clonación supone una **amplificación** genética. Ahora bien, a pesar de la definición anterior, entre los componentes de un clon también se dice que cada uno es un clon de los otros (por ej., cinco ovejas genéticamente idénticas forman un clon, y a la vez cada oveja es un clon de cualquiera de las demás, o todas son clónicas entre sí).

Debe indicarse que la clonación es un proceso fisiológico normal: son clónicas las *células embrionarias totipotentes* formadas por mitosis a partir del cigoto; cada una de estas células da lugar (mediante complejos procesos de diferenciación y desarrollo) a individuos adultos formados por células somáticas clónicas, genéticamente iguales entre sí y a la célula embrionaria de partida (excepto en lo que afecta a su *diversidad o polimorfismo genético*, resultante de mutación, recombinación, transposición, etc.). En el nivel de organismos completos, son clónicos los gemelos monocigóticos o univitelinos, procedentes de un mismo óvulo fecundado.

El interés de la clonación radica esencialmente en su carácter biotecnológico como proceso de multiplicación de moléculas de DNA o RNA, células, tejidos u organismos completos, mediante mecanismos de manipulación genética. La Biología en general, y la investigación biomédica en particular, están avanzando enormemente y lo harán aún más en el futuro, con la variada gama de copias de *elementos genéticos* ya empleados para estudiar y resolver problemas tanto de carácter básico como aplicado.

Globalmente, interesa describir la clonación desde el punto de vista de su finalidad, es decir:

- **Clonación de organismos:** plantas o animales completos.
- **Clonación de células,** aisladas o formando tejidos u órganos.
- **Clonación de moléculas,** sean éstas genes o fragmentos de DNA o RNA. De acuerdo con el proceso experimental, esta clonación puede dividirse a su vez en dos tipos:
 - **Clonación acelular,** es decir, clonación de moléculas sin intervención de célula alguna. Es sinónimo de *amplificación in vitro* de moléculas de DNA o RNA mediante la **reacción en cadena de la DNA polimerasa (PCR, polymerase chain reaction)**. Se describe en este tema.
 - **Clonación celular** (con células, no de células). Es la amplificación de moléculas de ácido nucleico mediante su introducción, asistida por vectores, en células *anfitrionas* en cultivo. Emplea la **tecnología del DNA recombinante**. Se estudia en el tema 16.

15.2 AMPLIFICACIÓN IN VITRO DEL DNA: REACCIÓN EN CADENA DE LA POLIMERASA O PCR

Considerada hoy día como una herramienta imprescindible en el laboratorio de Biología Molecular e Ingeniería Genética, el objetivo de esta técnica es la amplificación directa de un gen o un fragmento de DNA, o indirecta de un RNA (en este caso, a través de su DNA complementario, o cDNA), presentes en mezclas de muy diversas fuentes, sin necesidad de una purificación previa de la muestra integral original. Se puede partir de homogenados, extractos crudos de tejido, sangre completa, mezclas de fragmentos de DNA obtenidos con enzimas de restricción, muestras resultantes de la extracción y aislamiento de DNA, etc. Sin embargo, debe resaltarse que, por razones que se comprenderán más adelante, es un requisito imprescindible que **se conozca la secuencia de una parte** de la región de DNA o RNA que se quiere amplificar (salvo en el caso de algunas variantes del método clásico, como la PCR inversa, pág. 194).

Por estas características, en especial la capacidad de amplificación, la PCR es un método muy adecuado para preparar ácidos nucleicos en una cantidad muy superior a la de la muestra original, tanto como método de clonación acelular como para la detección.

Objetivo básico de la PCR: amplificar DNA	AMPLIFICACIÓN propiamente dicha Para disponer de cantidad suficiente como para utilizarlo, con fines diversos
	DETECCIÓN Para poderlo detectar en muestras con pequeñas cantidades del DNA diana

Las **aplicaciones** de la PCR son muy numerosas y variadas. Como ejemplos, pueden citarse: clonación acelular de fragmentos de DNA, detección de secuencias sin purificación previa, secuenciación de ácidos nucleicos, establecimiento de polimorfismos de secuencia, rastreo de mutaciones, **tipado de DNA** para trasplantes, diagnóstico de enfermedades genéticas, prenatales o no, determinación de secuencias específicas de DNA relacionadas con situaciones patológicas definidas, resolución de problemas forenses o arqueológicos, estudios evolutivos, detección de microorganismos infecciosos, detección de células tumorales, amplificación de DNA para su posterior clonación celular, etc.

15.2.1 Principio del método

La PCR no es una técnica analítica *per se*, sino más bien una metodología, resultante de la aplicación práctica de tres conceptos, ya descritos (págs. 163, 167 y 141, respectivamente):

1. **Desnaturalización** del DNA para dar hebras sencillas.
2. **Hibridación** específica de la hebra sencilla con un oligonucleótido. También se denomina etapa de **templado** (entendido como disminución de temperatura que permite la reasociación de hebras sencillas tras la

desnaturalización térmica; en inglés, *annealing*). Ésta es la etapa que explica por qué se debe conocer parte de la secuencia.

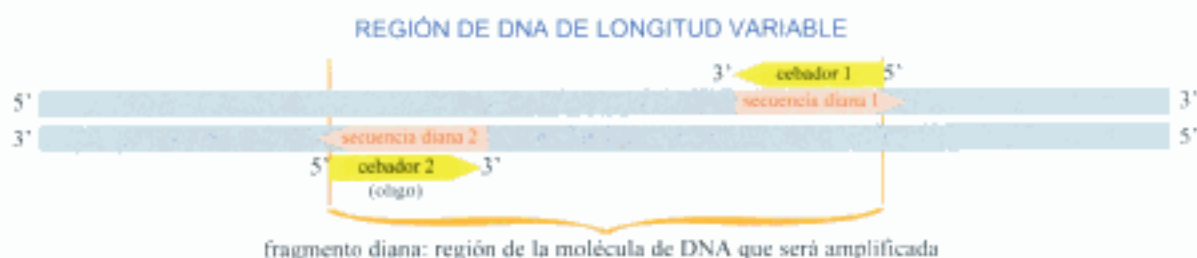
3. **Replicación** de la hebra sencilla por una DNA polimerasa a partir del oligonucleótido anterior como cebador: también, **elongación** o **extensión** del cebador, o **polimerización**.

Mediante la aplicación en forma cíclica de estos tres procesos se consiguen múltiples copias del fragmento de ácido nucleico en un corto espacio de tiempo, como se verá a continuación.

A pesar de esta aparente complejidad, en teoría es un método sencillo, sensible y relativamente rápido para amplificar secuencias. En la práctica, se requiere un control preciso de las variables que condicionan este triple proceso (secuencia diana, cebador, enzimas y resto de componentes), además de instrumentos adecuados (*termocicladores de DNA*) para establecer las condiciones de cada etapa y repetirlas cíclicamente (tiempo, temperatura, número de ciclos, etc.). Se consigue así amplificar secuencias de tamaños diversos, comprendidos entre 50 pb y 2.5 kb.

Para la realización práctica se precisan en la mezcla de reacción los siguientes "reactivos":

- a) Los cuatro **dNTPs**, en exceso, como sustratos para la síntesis de las innumerables copias de DNA. Para ser reconocidos por la polimerasa, deben ir acompañados de Mg^{2+} .
- b) Dos **oligonucleótidos** de cadena sencilla (coloquialmente, "oligos"), generalmente sintéticos, de 18 a 30 nt. Sus secuencias han de ser complementarias, respectivamente, a los dos extremos 3' de la región diana, uno en cada hebra, de modo que los oligos puedan ejercer de **cebadores** para la replicación de las dos hebras en la región diana. Por esta razón no se puede amplificar una región de DNA si no se conoce la secuencia de sus dos extremos. De hecho, es la posición donde hibridan los cebadores la que define la longitud del fragmento que se amplifica (secuencia diana).



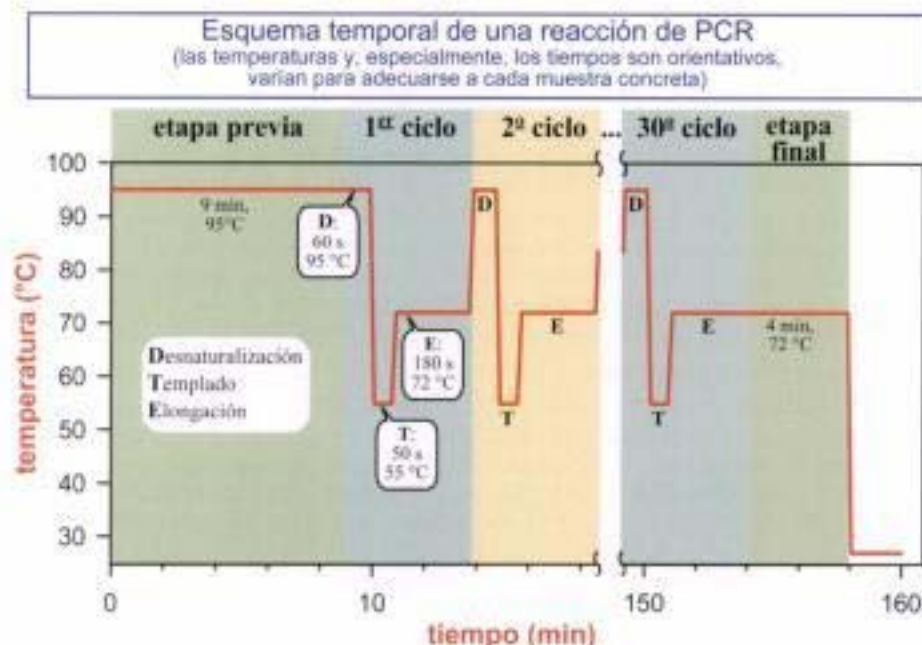
- c) Una **DNA polimerasa termoestable**, es decir, enzimáticamente activa a temperaturas relativamente altas (al menos 75 °C, y estable con frecuencia hasta 95 °C). Esta característica permite su actuación en sucesivos ciclos sin inactivarse; además, la replicación a temperaturas elevadas impide la formación de híbridos parcialmente desapareados y contribuye a la especificidad y rendimiento del proceso. La enzima más empleada se llama *polimerasa Taq*, por proceder de la bacteria *Thermus aquaticus*, que vive en manantiales de agua caliente. Esta enzima tiene el inconveniente de carecer de la actividad correctora de pruebas (exonucleasa 3'), por lo que la frecuencia de errores es superior a la de la replicación (alrededor de un error por cada 5.000 nucleótidos incorporados); de ahí que en ocasiones se sustituya por otras DNA polimerasas termoestables.

15.2.2 Etapas del proceso

Como ya se ha indicado, se requiere una sucesión de ciclos (generalmente entre 20 y 40, de 1.5 a 5 minutos de duración cada uno) de desnaturalización, hibridación y replicación, para conseguir una enorme amplificación del número de moléculas que contienen la secuencia de interés o diana. La duración total es de alrededor de 2 horas, dependiendo de las condiciones experimentales concretas.

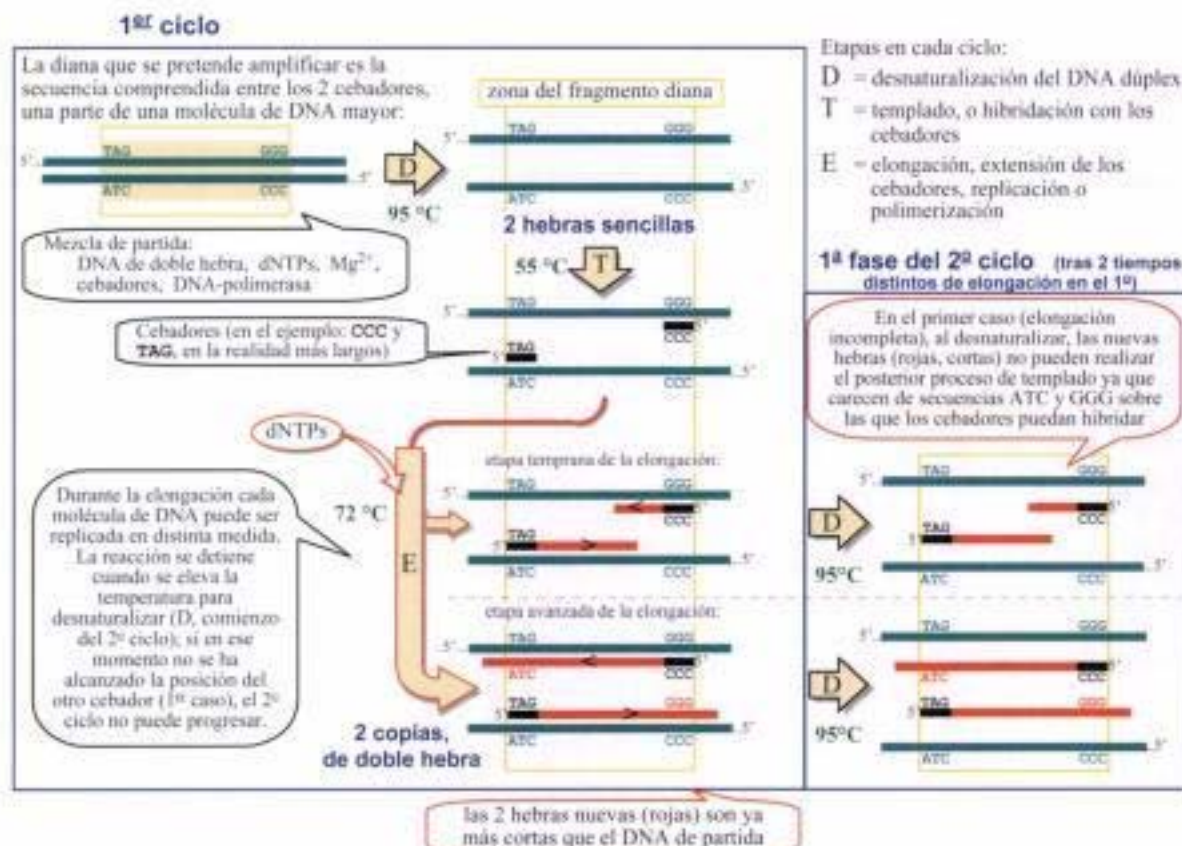
Cada ciclo de una PCR consta, pues, de tres etapas:

1. **Desnaturalización**: Calentamiento para la separación de las dos hebras del DNA, mediante una incubación breve (30-120 s) a una temperatura entre 68 y 97 °C, que debe ser superior a la de fusión (T_m , pág. 166) de la región de DNA que se quiere amplificar.
2. **Templado** (o hibridación): Enfriamiento rápido por debajo de T_m , de forma que se permite la hibridación (pág. 167) de las hebras sencillas del DNA de interés con los oligos cebadores. Generalmente se usan temperaturas de 37 a 65 °C que se mantienen entre 10 y 120 s.
3. **Elongación** (o replicación): Etapa de amplificación propiamente dicha (72-75 °C, 1-3 min), en la que la DNA polimerasa termoestable elonga los cebadores, empleando como molde ambas hebras originales. La replicación transcurre (pág. 146) en dirección 5'→3' a partir del extremo 3'-OH de cada cebador, empleando como sustrato los cuatro dNTPs, hasta terminar la lectura del molde o hasta que se comience una nueva etapa de desnaturalización (siguiente ciclo).

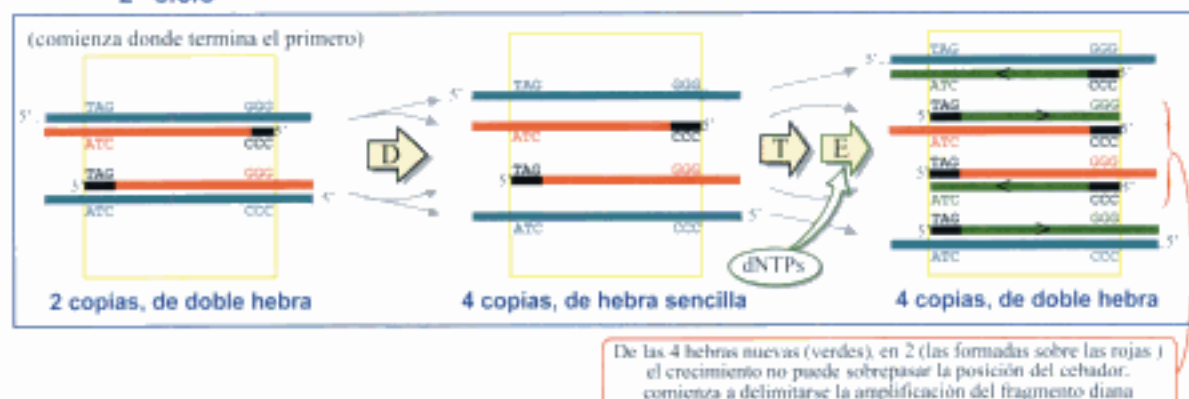


Además de las 3 etapas de cada ciclo, comúnmente se añade una etapa previa y una final al conjunto de todos los ciclos. La previa, a elevada temperatura (incluso superior a la de las etapas de desnaturalización), sirve básicamente para inactivar proteasas y nucleasas de la muestra, así como para asegurar la desnaturalización completa del DNA de partida, especialmente si éste es de gran tamaño o posee regiones muy compactadas (caso de un DNA genómico completo). La etapa final, por su parte, consiste en una prolongación de la última elongación, para permitir que se completen todos los fragmentos.

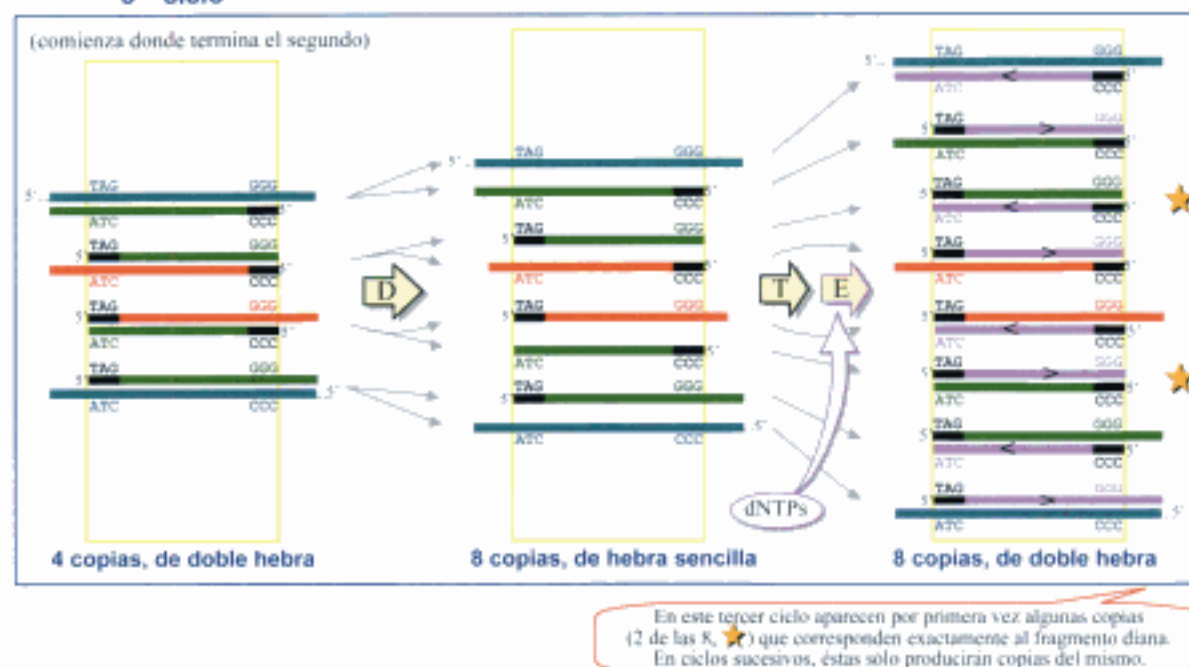
Esquemáticamente, el conjunto de reacciones y los resultados correspondientes a los 4 primeros ciclos son los siguientes:



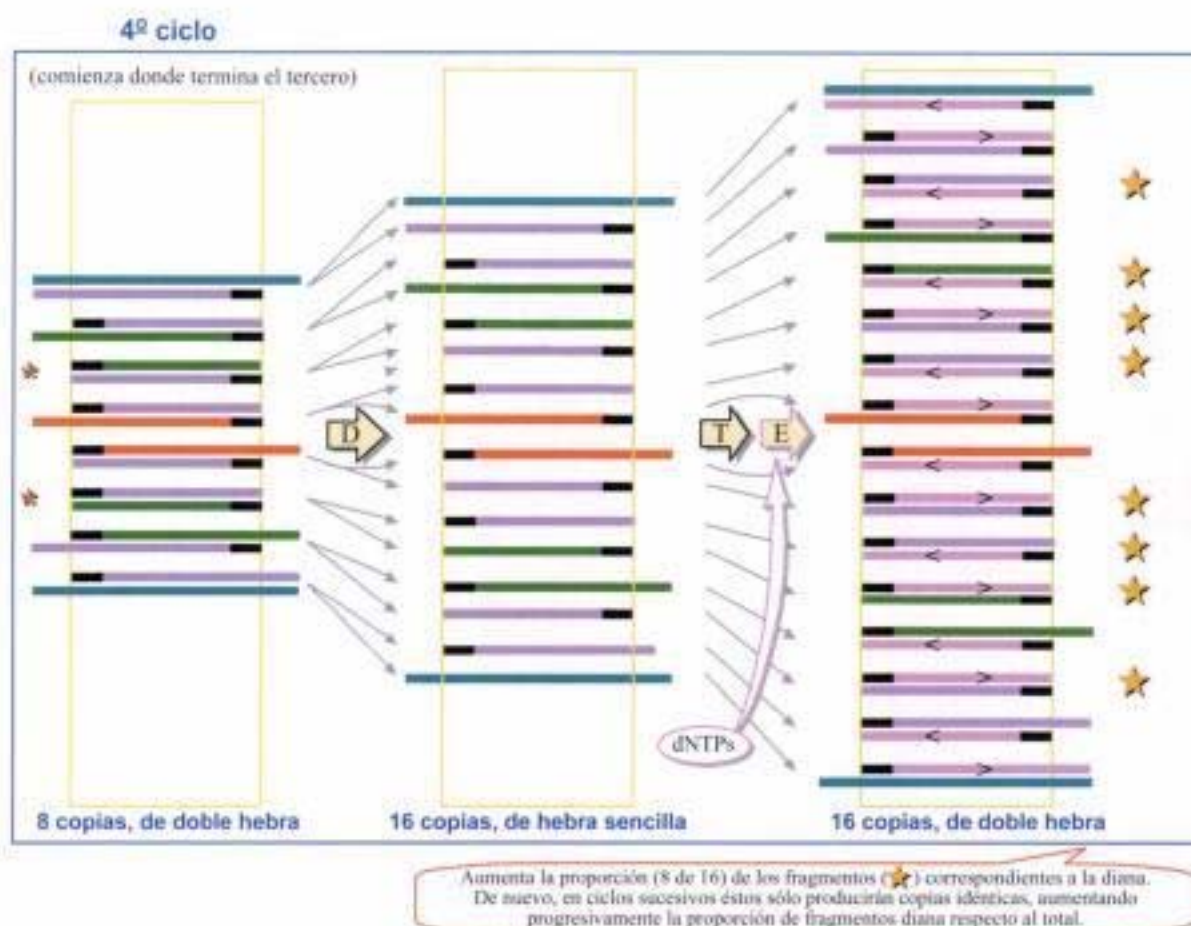
2º ciclo



3er ciclo



Debe hacerse notar que si la muestra de partida, como ocurre en el ejemplo descrito, está formada por moléculas de DNA de mayor tamaño que la región diana sólo a partir del tercer ciclo aparecen copias exactas de la diana. Si la muestra de partida hubiese sido el propio fragmento diana, todas las copias aparecidas desde el primer ciclo serían iguales a él.



Además de la amplificación global, por aumento del número de moléculas de DNA, debe notarse cómo el número de copias de la diana aumenta mucho más rápidamente que el de otras copias más largas. Todo ello hace que en la mezcla final, al terminar la PCR, los fragmentos diana superen abrumadoramente en número a cualquier otro tipo de fragmento, incluyendo obviamente las moléculas de DNA originales de la muestra (que no han sido amplificadas). Obsérvese en la tabla siguiente el enorme grado de amplificación conseguido, de un millón de veces en 20 ciclos y un billón en 40 ciclos. Esto es lo que permite aplicar la PCR a muestras poco purificadas y analizar los resultados directamente, en los que sólo se podrá detectar el fragmento amplificado.

nº de ciclos realizados	nº total de copias	nº de copias "largas"	nº de copias diana	dianas/total
1	2	2	0	0
2	4	4	0	0
3	8	6	2	25%
4	16	8	8	50%
5	32	10	22	69%
10	1.024	20	1.004	98%
20	$\approx 10^6$	40	$\approx 10^6$	99,996%
40	$\approx 10^{12}$	80	$\approx 10^{12}$	100%
x	2^x	2x	$2^x - 2x$	

15.2.3 Algunas características del proceso

Rendimiento: Puesto que los productos de cada ciclo sirven de molde para el siguiente, la acumulación de copias (tanto totales como dianas) es exponencial en vez de lineal. Sin embargo, en la práctica el rendimiento de cada etapa no es completo, por lo que las cifras se ven algo reducidas. A pesar de ello, siguen siendo muy elevadas. El número de copias al final de los ciclos realizados se puede calcular como:

$$N = N_0 (1 + R)^X$$

siendo: N = número de copias obtenidas
 N_0 = número de moléculas iniciales
 R = rendimiento de cada ciclo (generalmente, de 70 a 85%), expresado en tanto por uno (0,70-0,85)
 X = número de ciclos

Se han desarrollado modificaciones del método descrito, encaminadas a aumentar la eficiencia, sobre todo en lo relativo a la concentración de Mg^{2+} (normalmente de 1 a 4 mM) y al tiempo y temperatura de templado.

Duración: La duración de cada una de las tres etapas de cada ciclo y el número de ciclos deben optimizarse, dependiendo de la secuencia concreta que se quiera amplificar. En muchos casos, la elección de estas condiciones debe hacerse de forma empírica.

Especificidad: Es muy elevada. Está determinada especialmente por la secuencia de los dos cebadores utilizados y por las condiciones de templado. Se puede determinar la especificidad analizando por electroforesis los productos generados: la aparición de una banda única indica que la técnica es específica. Por ejemplo, si la temperatura de templado se reduce excesivamente se producen falsos positivos debidos fundamentalmente a apareamientos imperfectos del cebador con secuencias distintas de la diana, aunque similares, que resultan así también amplificadas.

Capacidad de detección (común, aunque incorrectamente, denominada **sensibilidad**): Es muy alta (bajo límite de detección), tanto que la PCR puede permitir detectar una única molécula de DNA en casi cualquier tipo de muestra clínica, forense o arqueológica (célula, pelo, semen, etc.). Sin embargo, esta característica supone también un elevado riesgo de contaminación por moléculas de origen ajeno a la muestra.

Fidelidad: Es decir, la capacidad de copiar secuencias con precisión, sin introducir mutaciones (cambios en la secuencia de nucleótidos). Esta característica depende sobre todo de que la enzima empleada tenga o no actividad correctora de pruebas (3'-exonucleasa, pág. 147). Así, la frecuencia de error oscila entre $2 \cdot 10^{-4}$ de una polimerasa *Taq* convencional y 10^{-6} de la polimerasa *Pfu*, que sí tiene actividad 3'-exonucleasa. En todo caso, la elección de la enzima dependerá del objetivo de la PCR. No es muy importante la fidelidad si sólo se pretende obtener una sonda, pero es fundamental si el objetivo es la identificación de mutaciones poco frecuentes.

15.2.4 Automatización de la PCR

La automatización es sencilla y de bajo coste, dado el carácter cíclico del proceso, el empleo de componentes estables al calor y el desarrollo de termocicladores para programar los cambios de temperatura. Al mismo tiempo es imprescindible, pues sólo mediante la automatización se hace factible la realización de un número elevado de ciclos y se aumenta la precisión del proceso. Ello facilita un empleo cada vez más específico y sensible en el laboratorio clínico.

15.2.5 Variantes del método

Han surgido numerosas modificaciones derivadas del método básico inicial, con el propósito de mejorar el rendimiento o la especificidad, adaptarse a muestras particulares, amplificar moléculas de RNA en lugar de DNA, producir cadenas de hebra sencilla... Por ejemplo, una modificación muy común en la actualidad (denominada "**comienzo en caliente**", *hot start PCR*) consiste en privar a la mezcla de reacción de alguno de sus componentes (frecuentemente la polimerasa) hasta que se haya alcanzado una temperatura superior a la de templado. De este modo, se evita la elongación de cebadores asociados inicialmente con poco rigor (a secuencias de homología parcial con la diana) y, como resultado, se aumenta mucho la especificidad de la amplificación.

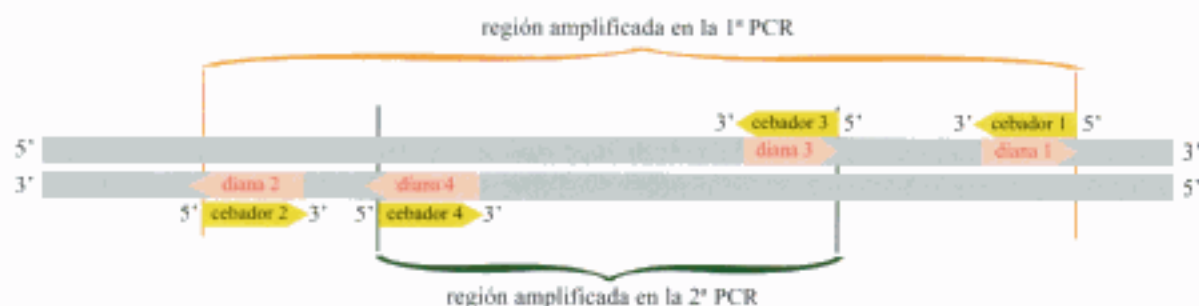
A continuación, se comentan brevemente algunas de las variantes más significativas.

15.2.5.1 PCR "larga"

Denominada **L-PCR** (*Long PCR*) o **LA-PCR** (*Longer and Accurate PCR*, "PCR más larga y exacta"), su objetivo es superar los límites de la PCR convencional para amplificar con fidelidad regiones diana de gran tamaño (entre 5 y 40 kb). El principal factor limitante es la ausencia de actividad correctora de pruebas en la polimerasa *Taq* (pág. 189), por lo que se añade una cantidad menor de otra enzima con capacidad de corrección de pruebas (por ejemplo, la *Pfu*) para contrarrestar la carencia de esta actividad y, al mismo tiempo, seguir aprovechando la eficacia de elongación de la polimerasa principal (*Taq* o similar).

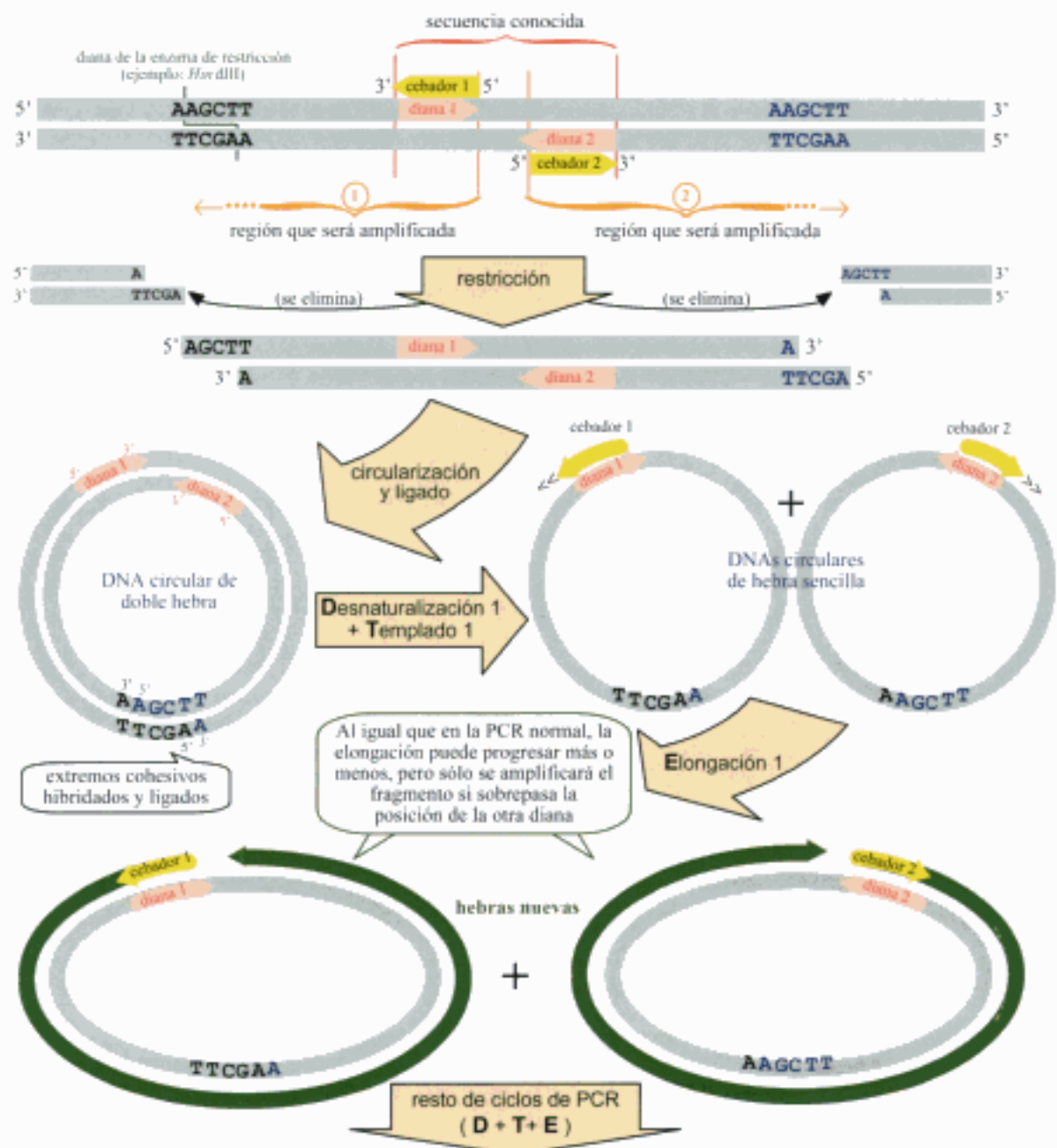
15.2.5.2 PCR "anidada"

La especificidad puede aumentarse realizando una segunda reacción de PCR, con dos cebadores nuevos que hibridan dentro del fragmento diana amplificado por el primer par, dando así lugar a productos de PCR más cortos, pero más específicos. Las copias correctas de la primera diana contendrán ambas secuencias complementarias a los nuevos cebadores, a diferencia de los productos no específicos generados en la primera PCR (que se observan en la electroforesis como varias bandas o una banda difusa), que por ello no resultarán amplificados en la segunda. Este método también sirve para aumentar el factor de amplificación conseguido, al suponer dos rondas de amplificación.



15.2.5.3 PCR inversa

Esta variante se emplea para clonar regiones desconocidas de un DNA, situadas en posición vecinal a secuencias diana conocidas. Es decir, en lugar de amplificar la región interna, *flanqueada* por los dos cebadores (PCR convencional), se amplifica la región externa, *que flanquea a* los cebadores. Para ello es necesario cortar el DNA a ambos lados de la región diana con una enzima de restricción (pág. 197), de tal forma que los extremos cohesivos resultantes puedan hibridar entre sí, formando una molécula circular. Ésta es cerrada por una ligasa y se realiza la PCR

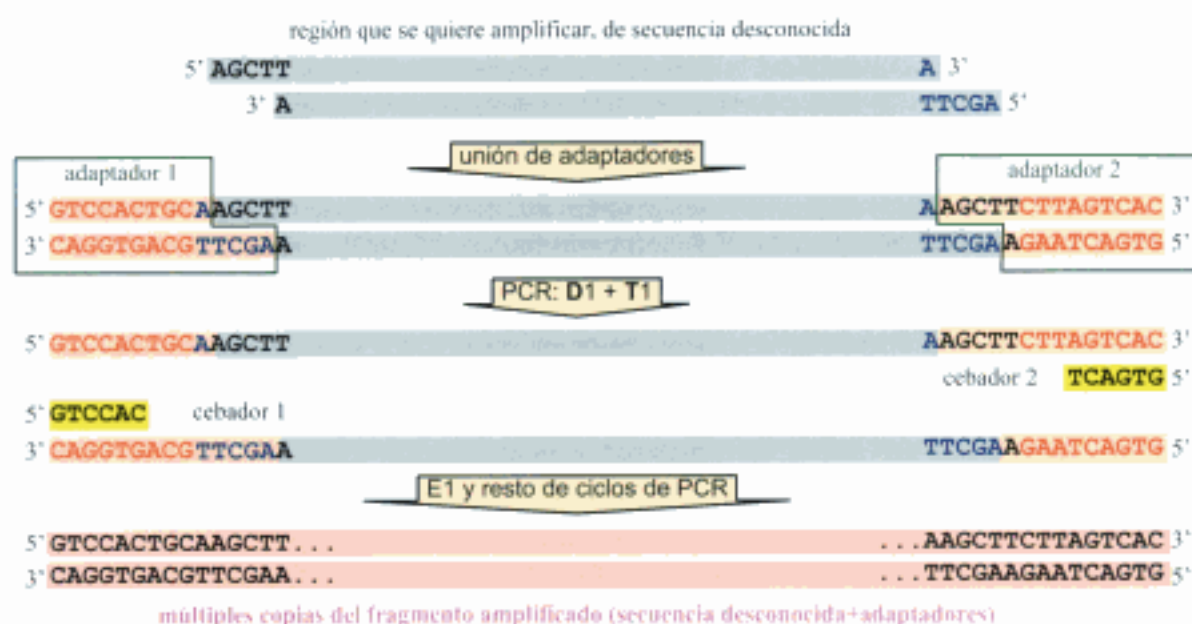




con cebadores que hibridan con los **extremos 5'** de la secuencia conocida, por lo que la elongación se extenderá alrededor del círculo. Se generan copias de un DNA lineal delimitado, como en la DNA normal, por la posición de ambos cebadores.

15.2.5.4 PCR con adaptadores

También puede amplificarse una región de DNA de secuencia desconocida ligando a sus fragmentos de restricción unos **adaptadores**, oligonucleótidos sintéticos con extremos cohesivos compatibles con los generados en la muestra. A continuación, una vez desnaturalizados, se añaden cebadores específicos para las secuencias 3' de los adaptadores. Se amplifica así el conjunto de adaptadores y secuencia diana, pero debe señalarse que se amplifican por igual todos los fragmentos de restricción presentes, no uno sólo.



15.2.5.5 PCR asimétrica

En este caso, se trata de generar **copias de hebra sencilla** de un DNA. En la variante más simple, se añaden cantidades muy diferentes de ambos cebadores, de modo que tras los primeros ciclos de PCR uno de ellos se agota y, disponibles ya suficientes copias del DNA diana, sólo una de sus hebras sigue amplificándose gracias al cebador más abundante.

15.2.5.6 RT-PCR: amplificación de RNA

El nombre, "PCR con transcriptasa inversa" (iniciales de *reverse transcriptase*), indica que se trata de una amplificación de RNA (especialmente mRNA) a través de la síntesis previa de su **cDNA** (DNA complementario al RNA), que después se amplifica por PCR. Es decir, no se obtienen copias del RNA de partida, sino de DNA, aunque se conserva obviamente la secuencia de aquél.

Es el método con mayor capacidad de detección de los disponibles para la medida de la expresión génica *in vitro*. Se utiliza preferentemente para amplificar mRNAs de células aseptibles, por ejemplo sanguíneas, con una expresión

elevada de ciertos genes (por ejemplo, globina). También se ha aplicado al análisis de transcritos de genes expresados en grado mínimo.

La mezcla inicial contiene todos los componentes necesarios: muestra de RNA, transcriptasa inversa, DNA polimerasa, cebadores y dNTPs. El proceso comienza (por ejemplo, 45 min a 48 °C) con la síntesis de una hebra de cDNA por la acción de la transcriptasa inversa, una polimerasa de DNA dirigida por RNA (pág. 147), permaneciendo el cDNA unido al molde como dúplex RNA : cDNA. En una segunda etapa (por ejemplo, 2 min a 94 °C) se desnaturaliza el dúplex y comienza la amplificación según el mecanismo de una PCR normal: la hebra de cDNA liberada actúa como molde para una segunda hebra de cDNA y luego el dúplex se amplifica en sucesivos ciclos.



Teóricamente, basta con una molécula de RNA que esté intacta entre los dos sitios de unión al cebador para conseguir la amplificación. Bajo el mismo principio, también se puede amplificar RNA usando la **polimerasa Tth** (también termoestable, aislada de la bacteria termófila *Thermus thermophilus*), que contiene al mismo tiempo actividades DNA polimerasa y transcriptasa inversa.

Tema 16

Clonación celular: Tecnología del DNA recombinante

16.1 Enzimas de restricción	197
16.1.1 Introducción a las nucleasas	198
16.1.2 Enzimas de restricción	199
16.1.2.1 Características generales	199
16.1.2.2 Papel biológico del sistema metilación-restricción	200
16.1.2.3 Acción de las enzimas de restricción de tipo II	201
16.2 Clonación celular de moléculas de DNA	203
16.2.1 Preparación del DNA para su clonación	205
16.2.1.1 Insertos de DNA	205
16.2.1.2 Insertos de DNA complementario procedente de un mRNA	205
16.2.2 Preparación del vector de clonación	206
16.2.3 Formación de la molécula de DNA recombinante	207
16.2.4 Tipos de vectores	208
16.2.4.1 Plásmidos	209
16.2.4.2 Bacteriófagos	210
16.2.4.3 Cósmidos	210
16.2.4.4 Cromosomas artificiales	211
16.2.5 Incorporación del DNA recombinante a la célula anfitriona	211
16.2.5.1 Tipos de célula anfitriona	211
16.2.5.2 Métodos de introducción del rDNA	211
16.2.6 Propagación en cultivo	213
16.2.7 Detección y selección de los clones recombinantes	213
16.2.8 Expresión génica	215
16.3 Genotecas	215
16.3.1 Genotecas genómicas	216
16.3.2 Genotecas de DNA complementario	216

El método clásico de clonación de ácidos nucleicos es el que aprovecha la capacidad de las células para replicar el DNA. Surge gracias a la aparición de una serie de técnicas que permiten el aislamiento del DNA y al descubrimiento en los años 70 de las enzimas de restricción, que hace posible en el laboratorio cortar el DNA e incorporar los fragmentos a vectores, que a su vez se introducen en células anfitrionas, donde tiene lugar su replicación. Aunque, en principio, esta tecnología sólo se planteaba para clonación del DNA, existen actualmente variantes que amplían su eficiencia y extienden su aplicación al RNA.

16.1 ENZIMAS DE RESTRICCIÓN

Antes de describir el proceso de clonación del DNA, es necesario hacer un inciso para estudiar con cierto detenimiento las propiedades y características de las **endonucleasas de restricción**. Aunque estas enzimas desempeñan una función propia en las células procariotas en las que se descubrieron, son particularmente importantes por su empleo experimental en varias áreas de la Biología Molecular, de interés tanto básico como aplicado al diagnóstico clínico. En especial, se utilizan en el proceso de la clonación molecular del DNA; de ahí la oportunidad de su estudio en este lugar. Además, tienen aplicación en otras técnicas, como la secuenciación del genoma y el estudio de

los polimorfismos (temas 18 y 27). Se inicia el estudio de estas enzimas con un breve recordatorio de las nucleasas en general.

16.1.1 Introducción a las nucleasas

Se denomina de forma genérica *nucleasa* a cualquier enzima con capacidad de escindir los enlaces fosfodiéster de la cadena polinucleotídica de los ácidos nucleicos; son, por tanto, fosfodiesterasas. Su especificidad puede analizarse con respecto a tres criterios:

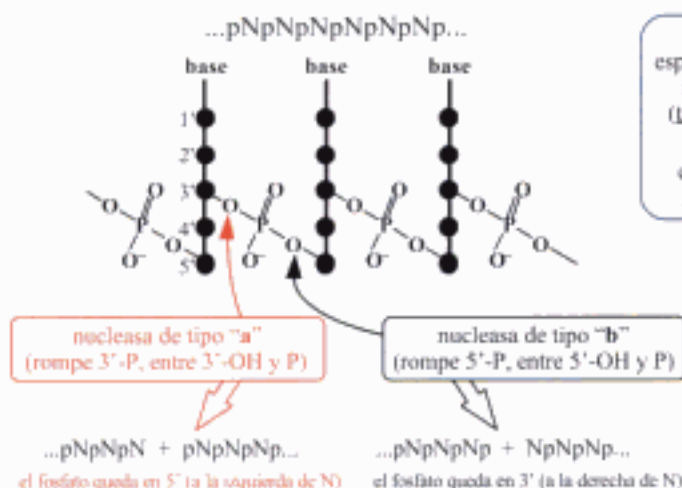
1. Escisión en **posiciones** internas o terminales en la estructura primaria (endonucleasas y exonucleasas, respectivamente):



Las **exonucleasas** hidrolizan exclusivamente enlaces fosfato en los que participa el nucleótido terminal (bien 5' o bien 3', dependiendo de la exonucleasa). Generan, por tanto, una cadena polinucleotídica acortada y un nucleósido (N), un nucleósido-monofosfato (pN o Np) o un nucleósido-bisfosfato (pNp).

Las **endonucleasas** sólo hidrolizan enlaces internos de la cadena, es decir, sin afectar a los nucleótidos terminales. Generan dos fragmentos polinucleotídicos.

2. Actuación sobre uno de los dos **tipos de enlace fosfoéster**:

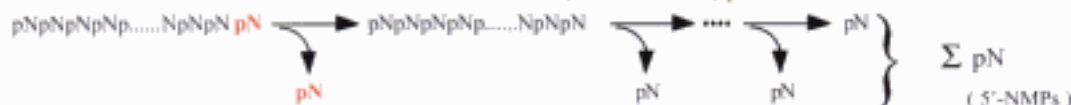


Algunas nucleasas (tipo "a") actúan específicamente hidrolizando los enlaces fosfato formados con el OH 3', mientras que otras (tipo "b") hidrolizan los formados con OH 5'.

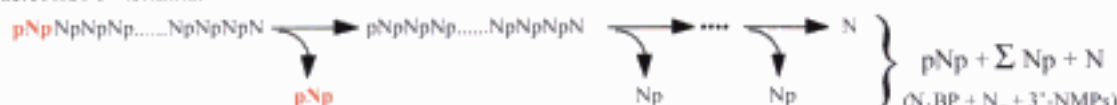
De este modo, las de tipo "a" dan lugar a extremos 3'-OH y 5'-P, mientras que las de tipo "b" producen extremos 3'-P y 5'-OH.

Ejemplos de especificidad tipo "a" y "b" de endonucleasas

Fosfodiesterasa I (de veneno de serpiente). Es una 3'-exonucleasa de tipo "a", por tanto libera el nucleótido 3'-terminal en forma de nucleósido-5'-P. Su acción exhaustiva (hidrólisis total) produce la mezcla de nucleósidos-5'-P:



Fosfodiesterasa II (de bazo bovino). Es una 5'-exonucleasa de tipo "b", por lo que libera el nucleótido 5'-terminal como mononucleósido-3',5'-bisfosfato. Si la escisión es repetitiva, aparecen los sucesivos mononucleósidos-3'-P y el nucleósido 3'-terminal:



3. Reconocimiento de nucleósidos, bien por la pentosa o bien por la base nitrogenada (sólo en algunas nucleasas): **desoxirribonucleasas** o DNasas, que actúan sólo sobre DNA; **ribonucleasas** o RNasas, que hidrolizan RNA; finalmente, algunas nucleasas distinguen entre nucleósidos purínicos y pirimidínicos:

Ejemplos de especificidad por nucleósidos

RNasa I (de páncreas bovino). Endorribonucleasa de tipo "b" que actúa sólo sobre enlaces en los que el ribonucleósido del lado 5' es pirimidínico (C,U). Se forman, por tanto, oligonucleótidos con extremo 5'-OH y extremo 3'-P pirimidínico (más, en su caso, el 5'-P-ribonucleósido pirimidínico 5'-terminal y el ribonucleósido 3'-terminal):



DNasa II (de hazo, timo, etc.). Endodesoxirribonucleasa de tipo "b" que hidroliza preferentemente el enlace entre nucleósidos purínico y pirimidínico; con menor frecuencia, entre pirimidínico y purínico o entre purínico y purínico. Se forman, pues, oligonucleótidos con extremos 5'-OH y 3'-P:



16.1.2 Enzimas de restricción

16.1.2.1 Características generales

Las enzimas de restricción se llaman también **nucleasas de restricción**, **endonucleasas de restricción** y, en ocasiones, **enzimas restrictivas** o **restrictasas**. Pertenecen a este grupo una gran diversidad de endodesoxirribonucleasas, descubiertas como enzimas propias de distintas bacterias y caracterizadas en los años 70. Se dispone hoy comercialmente de un gran número de ellas con elevada especificidad, estabilidad y pureza.

Se nombran con tres letras tomadas del género y especie de la bacteria de la que se aislaron originalmente, seguidas a veces por una letra más, que identifica el serotipo (variante antigénica de la bacteria), y finalmente por un número romano que las identifica en el caso de que en una misma variante se hayan encontrado varias enzimas con distinta especificidad:



La importancia de las enzimas de restricción radica en su gran **especificidad** de reconocimiento de una secuencia corta de DNA dúplex y la consiguiente hidrólisis de un enlace fosfodiéster en cada hebra, justamente en secuencias concretas del DNA llamadas **sitios de reconocimiento o de restricción**. Los fragmentos de DNA originados son útiles para iniciar la clonación celular o acelular, para el análisis de DNA, para elaborar mapas físicos de restricción (pág. 226), para la detección de polimorfismos (pág. 381), etc.

Las diversas enzimas de restricción se agrupan normalmente en tres familias, de acuerdo con sus propiedades:

Endonucleasas de restricción	Tipo I	Tipo II	Tipo III
Actividad enzimática	Endonucleasa y metilasa en una misma proteína (en distintas subunidades)	Endonucleasa (metilasa en una proteína independiente)	Endonucleasa y metilasa en una misma proteína (en distintas subunidades)
Coenzimas o cosustratos	ATP, Mg ²⁺ , SAM (*)	Mg ²⁺	ATP, Mg ²⁺ , SAM (*)
Estructura proteica	3 subunidades: S reconocimiento, R restricción, M metilación	1 subunidad, aunque actúa como homodímero	2 subunidades: MS reconocimiento y metilación, R restricción
Sitio de reconocimiento	No palindrómico, dos partes separadas (ej.: TGA(N) ₆ TGCT)	Palindrómico (4-8 pb) (pág. 57)	No palindrómico
Punto de corte	Corta las 2 hebras en puntos cercanos entre sí, poco específicos, a unos 1.000 pb del sitio de reconocimiento	Corta las 2 hebras en puntos equivalentes, muy específicos, dentro de la secuencia de reconocimiento	Corta las dos hebras en puntos distantes entre sí 2 ó 3 nt. situados a 24-26 pb del sitio de reconocimiento
Punto de metilación	Dentro de la secuencia de reconocimiento. En ambas hebras	Adenina o citosina muy específicas dentro de la secuencia de reconocimiento. En ambas hebras	Adenina dentro de la secuencia de reconocimiento. En una sola hebra
Interés en ingeniería genética	No	Si	No
Ejemplos	<i>Eco</i> K, <i>Eco</i> B ...	Numerosos (pág. 202)	<i>Eco</i> PI, <i>Eco</i> P15, <i>Hinf</i> III ...

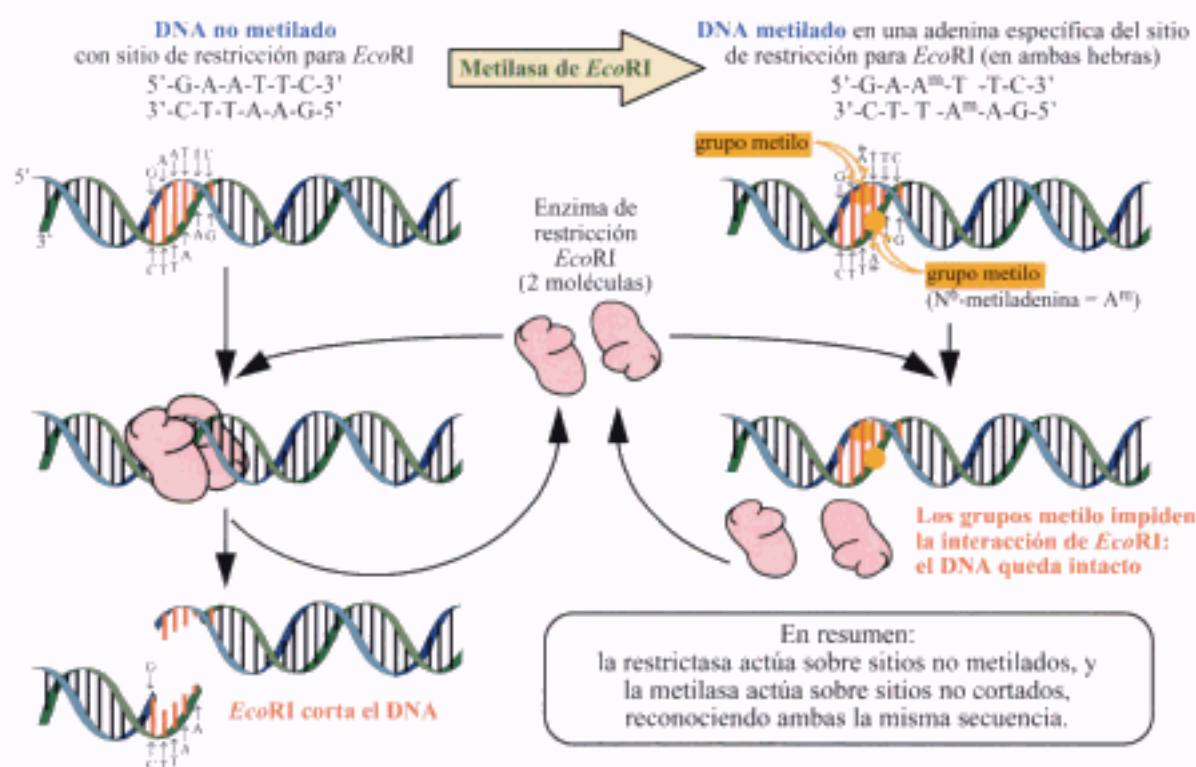
(*) SAM = S-adenosilmetionina (pág. 201)

Nos concentraremos en el estudio de las enzimas de tipo II, pues, al cortar en secuencias muy definidas, son las utilizadas como herramientas en ingeniería genética.

16.1.2.2 Papel biológico del sistema metilación-restricción

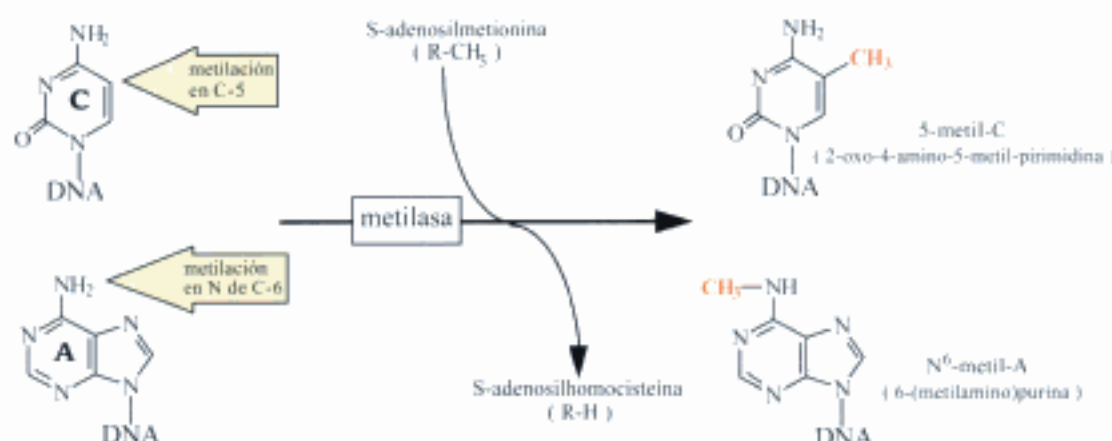
La existencia natural de las nucleasas de restricción en muchas bacterias ofrece a éstas un mecanismo de defensa contra la entrada de material genético de otro organismo, concretamente de virus bacteriófagos (cuya reproducción depende de la maquinaria genética de la bacteria): se trata de los sistemas de **restricción-modificación** o **metilación-restricción**. Para cada enzima de restricción, una metilasa reconoce la misma secuencia que constituye el sitio de restricción y une covalentemente grupos metilo a determinadas bases del DNA en dicha secuencia. En el caso de enzimas de tipo II, la metilasa es una proteína independiente, mientras que las de tipo I y III poseen las actividades nucleasa y metilasa en su molécula oligomérica, como subunidades que actúan coordinadamente (ver tabla anterior).

El mecanismo de defensa es el siguiente: el DNA propio de la bacteria es metilado de una forma específica, característica de cada especie bacteriana (en las secuencias reconocidas tanto por la restrictasa como por la metilasa de esa especie). La metilación de las bases impide la unión de la enzima de restricción, con lo que el DNA propio no es hidrolizado. Por el contrario, al entrar en la célula DNA de otro organismo, no metilado o con un patrón de metilación diferente, este DNA puede ser degradado por la enzima de restricción, ya que carece de los grupos metilo en la secuencia diana (por razones puramente aleatorias, en cualquier genoma habrá un cierto número de secuencias de 4, 6 u 8 pb iguales a la reconocida por la enzima, luego habrá secuencias diana disponibles). A partir de este mecanismo de acción combinada surgió precisamente el nombre de enzimas de restricción: la acción de la nucleasa *restringe* la posibilidad de infección por virus.



Si está metilada una sola hebra del DNA, el sitio no es reconocido por la enzima de restricción, pero sí por la metilasa, que añade el grupo metilo a la otra hebra. Esto es lo que ocurre, por ejemplo, al final de la replicación, donde el dúplex está formado por una hebra molde, preexistente y, en consecuencia, metilada, y otra hebra recién sintetizada, que aún carece de grupos metilo. Esta acción de las metilasas asegura que todo el DNA de las dos células hijas quede metilado lo antes posible.

La metilación del DNA afecta principalmente a citosinas y adeninas de la secuencia reconocida. En la reacción, las metilasas utilizan como donador del grupo metilo el compuesto S-adenosilmetionina (pág. 22).



Debe adelantarse que la metilación del DNA interviene en eucariotas como un mecanismo de regulación de la expresión génica, aunque este proceso no tiene nada que ver con el mecanismo de protección que desempeñan los sistemas metilasa-restrictasa en procariontes. Esencialmente, en eucariotas los genes tienden a estar desmetilados en los tejidos donde se expresan y metilados donde no se expresan, especialmente en las regiones de DNA llamadas "islotos CpG", por tener esta secuencia dinucleotídica en la que C sufre la metilación.

16.1.2.3 Acción de las enzimas de restricción de tipo II

Son éstas las restrictasas estructuralmente más simples y las mejor estudiadas por su utilidad en Ingeniería Genética, a modo de "tijeras" para el corte del DNA. El **sitio de restricción** es a la vez lugar de reconocimiento y de hidrólisis. Se hidrolizan de forma muy específica enlaces fosfoéster del DNA, uno en cada hebra, generándose dos **fragmentos de restricción**. El enlace hidrolizado es siempre el 3'-P (es decir, son endonucleasas del tipo "a", pág. 198), por lo que el fosfato queda unido a 5', dejando en ambos fragmentos extremos 3'-OH y 5'-P. La reacción no necesita ATP, lo que la diferencia de la catalizada por las enzimas de tipo I y III.

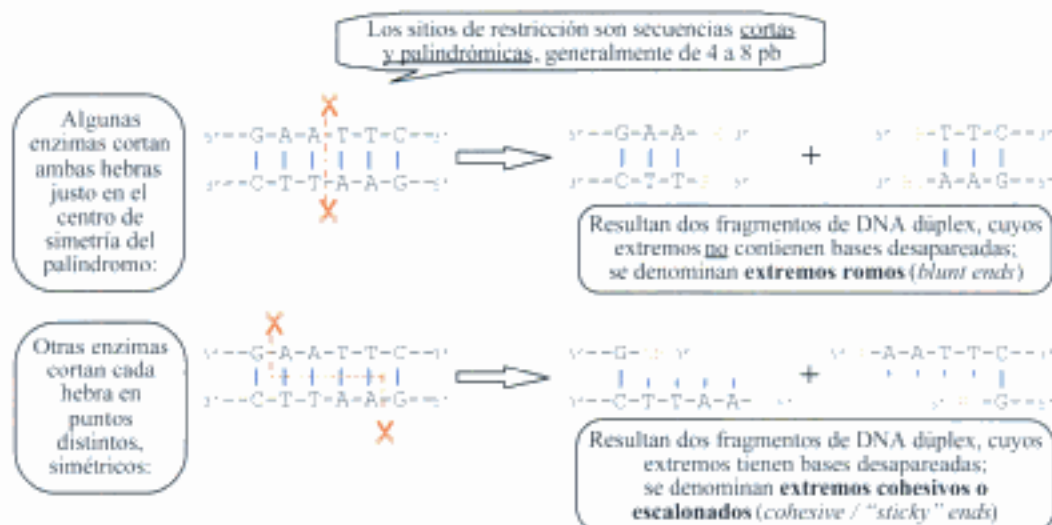
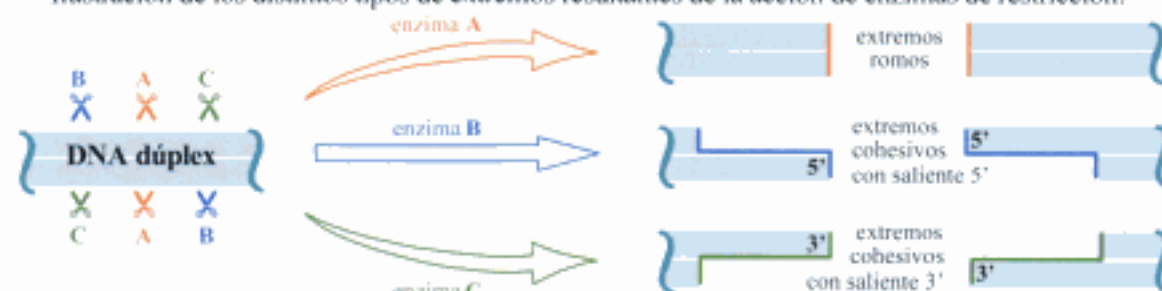


Ilustración de los distintos tipos de extremos resultantes de la acción de enzimas de restricción:



Algunas enzimas de restricción de tipo II empleadas en ingeniería genética

Nombre de la enzima	Bacteria de origen	Secuencia reconocida y puntos de corte (/) (N = cualquier nucleótido)	Tipo de extremos	Productos resultantes de la escisión
<i>Acc 65I</i>	<i>Acinetobacter calcoaceticus</i> 65	-G/G-T-A-C-C- -C-C-A-T-G-G-	cohesivos, saliente 5'	-G -C-C-A-T-Gp ★ pG-T-A-C-C- G-
<i>Afu I</i>	<i>Arthrobacter luteus</i>	-A-G/C-T- -T-C/G-A-	romos	-A-G -T-Cp pC-T- G-A-
<i>Bam HI</i>	<i>Bacillus amyloliquefaciens</i> H	-G/G-A-T-C-C- -C-C-T-A-G-G-	cohesivos, saliente 5'	-G -C-C-T-A-Gp ★ pG-A-T-C-C- G-
<i>Bgl II</i>	<i>Bacillus globigii</i>	-A/G-A-T-C-T- -T-C-T-A-G/A-	cohesivos, saliente 5'	-A -T-C-T-A-Gp ★ pG-A-T-C-T- A-
<i>Eco RI</i>	<i>Escherichia coli</i> RY13	-G/A-A-T-T-C- -C-T-T-A-A/G-	cohesivos, saliente 5'	-G -C-T-T-A-Ap pA-A-T-T-C- G-
<i>Hae III</i>	<i>Haemophilus aegyptius</i>	-G-G/C-C- -C-C/G-G-	romos	-G-G -C-Cp pC-C- G-G-
<i>Hin dIII</i>	<i>Haemophilus influenzae</i> Rd	-A/A-G-C-T-T- -T-T-C-G-A/A-	cohesivos, saliente 5'	-A -T-T-C-G-Ap pA-G-C-T-T- A-
<i>Kpn I</i>	<i>Klebsiella pneumoniae</i> OK8	-G-G-T-A-C/C- -C/C-A-T-G-G-	cohesivos, saliente 3'	-G-G-T-A-C -Cp ★ pC- C-A-T-G-G-
<i>Mbo I</i>	<i>Moraxella bovis</i>	-N/G-A-T-C-N- -N-C-T-A-G/N-	cohesivos, saliente 5'	-N -N-C-T-A-Gp ★★ pG-A-T-C-N- N-
<i>Nof I</i>	<i>Nocardia otitidis-caviarum</i>	-G-C/G-G-C-C-G-C- -C-G-C-C-G-G/C-G-	cohesivos, saliente 5'	-G-C -C-G-C-C-G-Gp pG-G-C-C-G-C- C-G-
<i>Pst I</i>	<i>Providencia stuartii</i>	-C-T-G-C-A/G- -G/A-C-G-T-C-	cohesivos, saliente 3'	-C-T-G-C-A -Gp pG- A-C-G-T-C-
<i>Pvu I</i>	<i>Proteus vulgaris</i>	-C-G-A-T/C-G- -G-C/T-A-G-C-	cohesivos, saliente 3'	-C-G-A-T -G-Cp pC-G- T-A-G-C-
<i>Pvu II</i>	<i>Proteus vulgaris</i>	-C-A-G/C-T-G- -G-T-C/G-A-C-	romos	-C-A-G -G-T-Cp pC-T-G- G-A-C-
<i>Sau 3AI</i>	<i>Staphylococcus aureus</i> 3A	-N/G-A-T-C-N- -N-C-T-A-G/N-	cohesivos, saliente 5'	-N -N-C-T-A-Gp ★ pG-A-T-C-N- N-
<i>Tth 111I</i>	<i>Thermus thermophilus</i> strain 111	-G-A-C-N/N-N-G-T-C- -C-T-G-N-N/N-C-A-G-	cohesivos, saliente 5'	-G-A-C-N -C-T-G-N-Np pN-N-G-T-C- N-C-A-G-

El nombre de la enzima es una abreviatura del de la bacteria en la que se encuentra

Las secuencias diana más frecuentes tienen 6 pb, pero también son comunes las de 4 pb y 8 pb (*Nor I*); se dan algunos casos con otro tamaño, por ejemplo, 9 pb (*Tth 111I*). Siempre son secuencias palindrómicas.

Tth 111I es un ejemplo de enzima con especificidad parcial de secuencia (acepta cualquier base en las tres posiciones N)

Extremos compatibles: en algunos casos, distintas enzimas de restricción generan extremos cuyo saliente tiene la misma secuencia. Ejemplo: *Bam HI*, *Bgl II* y *Mbo I* (★).

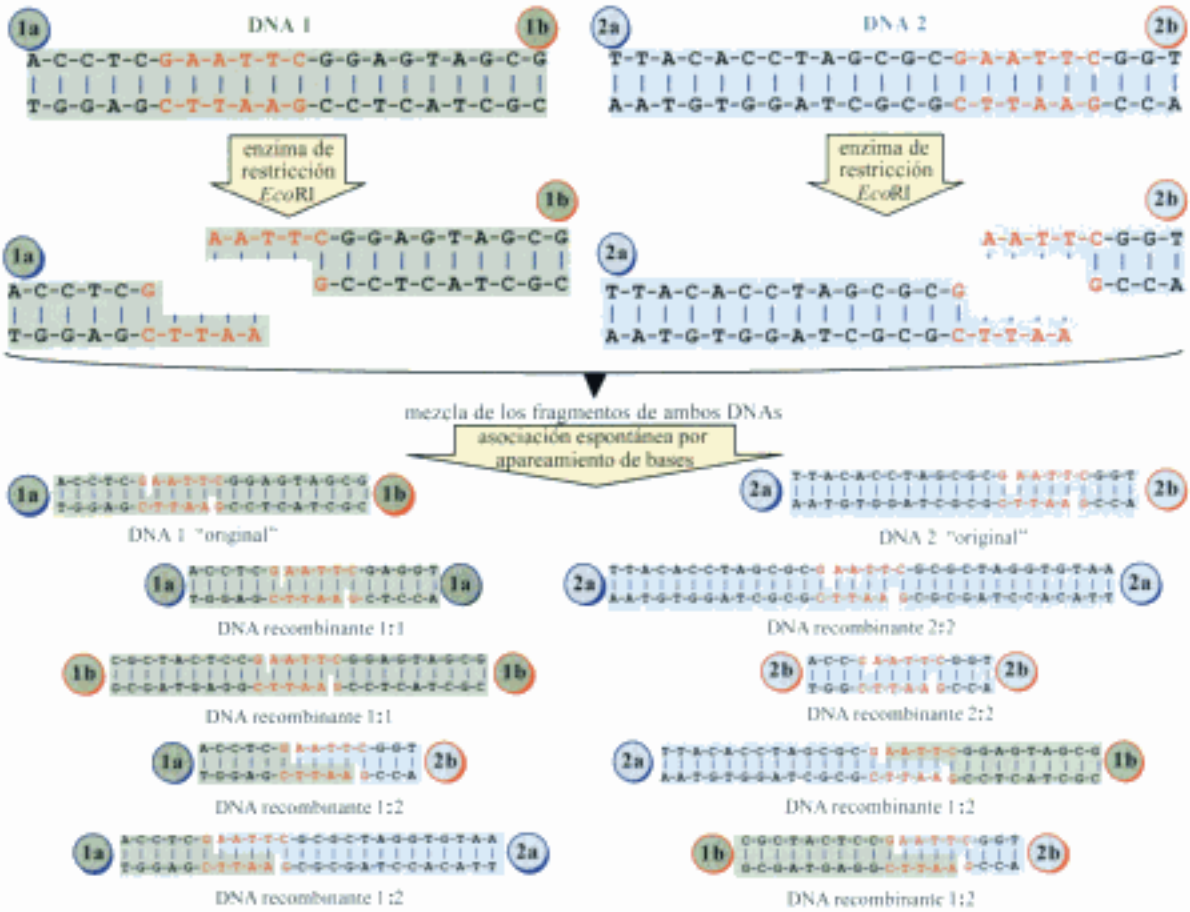
Isoesquizómeros nombre que se aplica a dos o más enzimas que reconocen la misma secuencia; pueden cortar en el mismo punto o bien se diferencian en que el corte o la metilación asociada tienen lugar en puntos distintos dentro de esa secuencia común de reconocimiento. Ejemplos: *Mbo I* y *Sau 3AI* (★); *Acc 65I* y *Kpn I* (★).

Obsérvese que en todos los casos los extremos de los dos fragmentos generados son idénticos, tienen la misma secuencia (se puede ver girando 180° uno de ellos). Esto es consecuencia de la simetría del palíndromo y de la posición simétrica de los puntos de corte en cada hebra. Todo ello se debe a que la enzima actúa en forma de dímero, por tanto reconociendo una misma secuencia y cortando un mismo tipo de enlace en las dos hebras: por ejemplo, *Eco RI* corta el enlace (5') G—AATTC (3') en ambas hebras.

La formación de extremos cohesivos es de gran interés aplicado en ingeniería genética (pág. 207), pues fragmentos generados con una misma restrictasa a partir de dos DNAs distintos pueden interaccionar mediante el apareamiento de las bases de sus extremos cohesivos (tal interacción no es posible entre extremos romos). Igualmente se pueden asociar los fragmentos producidos por enzimas distintas que producen extremos compatibles.

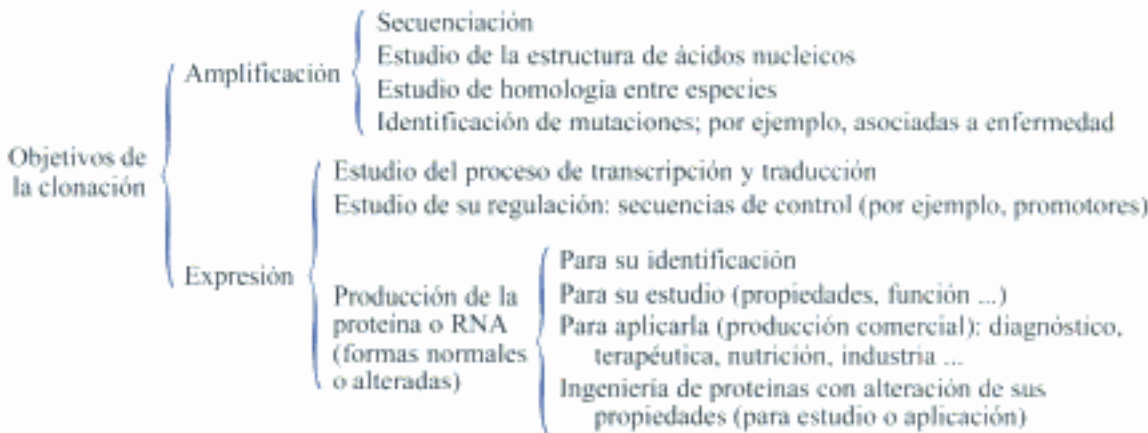
Debe resaltarse que el producto de la reasociación de los dos fragmentos en ninguno de los casos es idéntico al DNA de partida, pues está formado por dos moléculas. La falta de enlace covalente entre los extremos de los fragmentos reasociados se indica con el nombre de "mella" o "muesca" (nick, pág. 156). Éstas se pueden cerrar o "sellar" por la acción de las enzimas **ligasas** (o DNA ligasas), dando una molécula única de doble hebra. Como se ha visto (pág. 156), las ligasas catalizan la formación en cada hebra del enlace fosfoéster que faltaba, con consumo de energía. Pueden, incluso, unir o "ligar" dos fragmentos de DNA con extremos romos, aunque con mucha menor eficacia, pues no hay apareamiento entre ellos. La acción de las ligasas es, pues, contraria a la de las nucleasas (formación de enlace fosfoéster frente a su hidrólisis), pero el mecanismo de la reacción no es exactamente opuesto, por el

requerimiento energético de la primera. Debe resaltarse que la ligasa siempre requiere un grupo 5'-P para su actuación, propiedad que posee utilidad experimental (pág. 208).



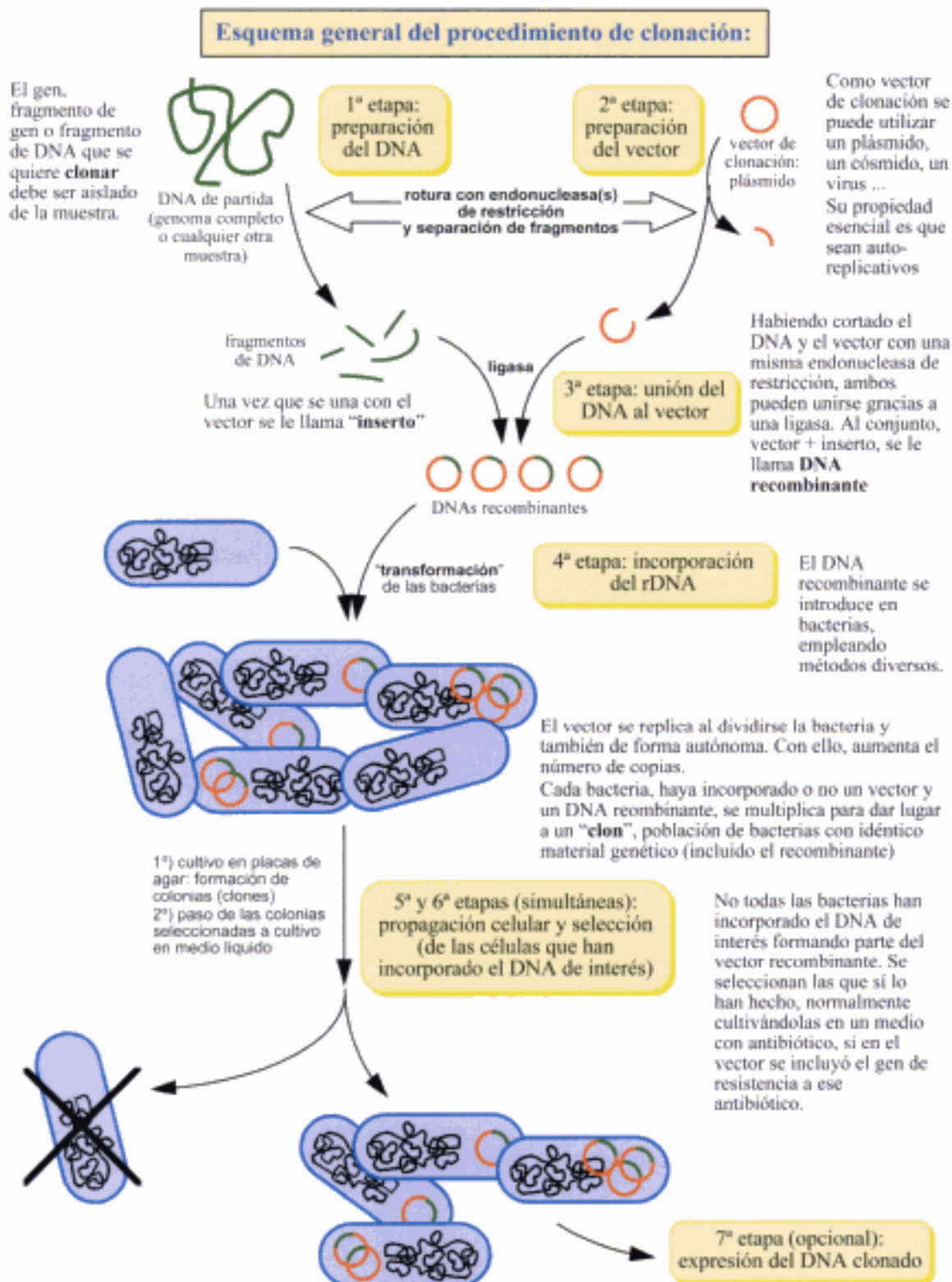
16.2 CLONACIÓN CELULAR DE MOLÉCULAS DE DNA

El objetivo central de esta clonación es la producción de un gran número de copias de una región de DNA (fragmentos o genes) o de cDNA (formado a partir de un mRNA). Se trata, por tanto, de un proceso de clonación de tipo **celular** (pues se consigue gracias al empleo de células, pág. 188) que, por contraposición a la clonación **acelular** por PCR, *in vitro*, se puede considerar que se realiza *in vivo* (estrictamente hablando, en cultivo). Se basa esta clonación, al igual que la PCR, en la realización de múltiples rondas de replicación, pero en este caso catalizadas por la DNA polimerasa propia de la célula en la que se lleva a cabo el proceso, célula **"anfitriona"**. Para ello, el fragmento de DNA o cDNA a clonar (llamado **inserto**, por razones que se comprenderán después) se une a otro DNA (**vector de clonación**), y la molécula de **DNA recombinante** formada se incorpora a la célula anfitriona donde tiene lugar la amplificación por replicación. Una vez detectada la presencia del DNA de interés y seleccionados los clones que lo han amplificado, se aísla el gen o fragmento de DNA o cDNA, que se emplea con distintos fines, esencialmente el estudio de características



físicas, secuencia, estructura, propiedades funcionales, expresión génica a proteínas, etc. (muchos de estos objetivos pueden ser aplicables también a la PCR o clonación acelular).

Por razones pedagógicas, a la vez que de enfoque experimental, puede describirse la clonación en **siete etapas**: dos de preparación de muestras, cuatro sucesivas de clonación propiamente dicha y, según los casos, una de expresión. En ocasiones, se añade la amplificación por PCR del DNA clonado (en este caso, el proceso comprendería dos clonaciones sucesivas, celular la primera y acelular la segunda). Hoy día, todas las etapas se realizan de forma rutinaria, a veces incluso automatizada, al disponerse de forma comercial de las enzimas, sustratos, vectores, etc., e instrumentos adecuados para su realización.

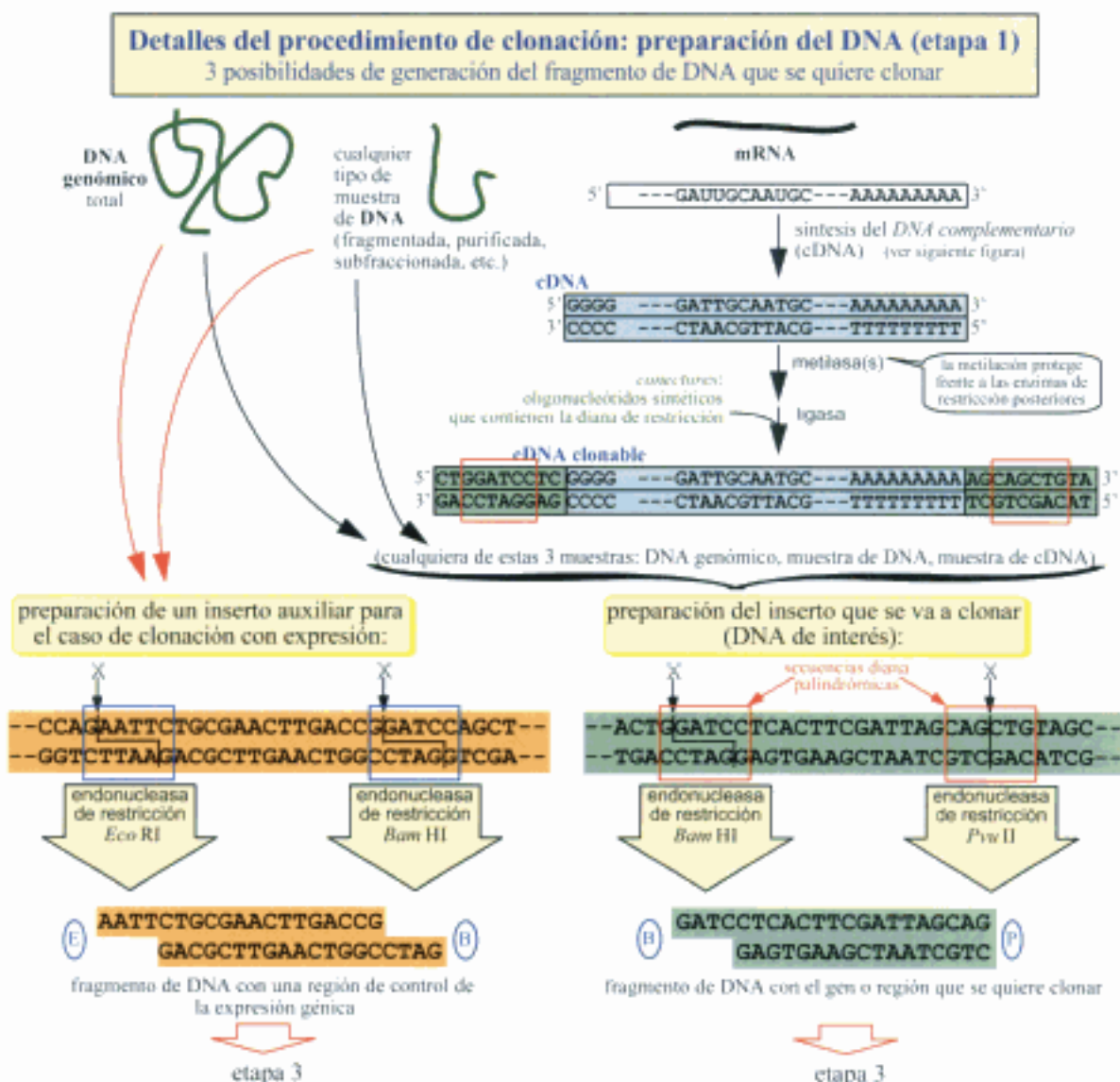


16.2.1 Preparación del DNA para su clonación

Esta primera etapa de la clonación, la preparación inicial del ácido nucleico que se quiere clonar, es una parte esencial del proceso.

16.2.1.1 Insertos de DNA

Se parte de células nucleadas, normalmente sanguíneas, lisadas por sonicación o por suspensión en un medio hipotónico. Una vez eliminadas las moléculas acompañantes, tanto RNA (con RNasas o hidrólisis alcalina) como proteínas (con proteasas) y enzimas o fragmentos proteolíticos (en sistemas bifásicos fenol/agua), se precipita y concentra el DNA (con etanol y sales). Éste posteriormente se **fragmenta** con enzimas de restricción, para dar lugar a una mezcla cuyo contenido, en número y tamaño de fragmentos, depende del número de sitios de restricción y de su accesibilidad para las enzimas. El fragmento que se va a clonar debe aislarse, o separarse del resto, mediante alguna técnica adecuada, por ejemplo electroforesis en gel de agarosa, centrifugación en gradiente de densidad, cromatografía líquida de alta presión (HPLC), etc. Cuando se desee expresar el fragmento clonado (etapa 7), es necesario preparara además un inserto auxiliar que contenga elementos reguladores de la expresión génica.



16.2.1.2 Insertos de DNA complementario procedente de un mRNA

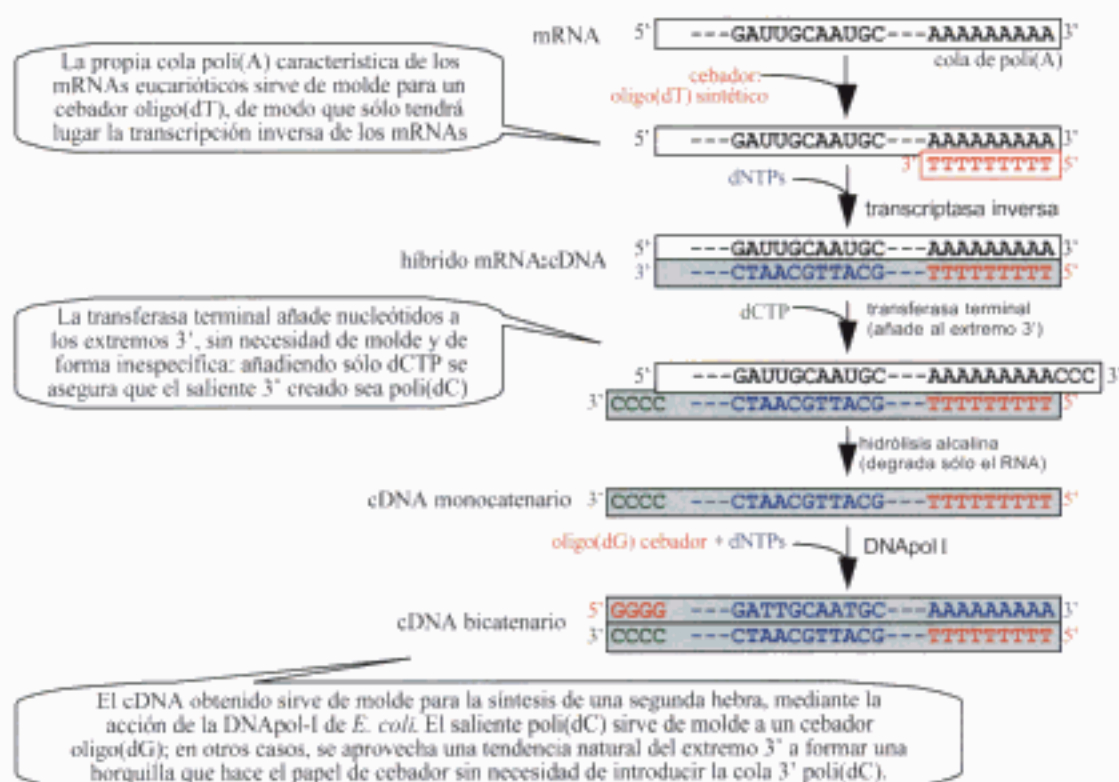
En muchos casos interesa clonar la secuencia codificante de un gen eucariota, para lo cual se parte de muestras de mRNA que debe ser convertido en el cDNA correspondiente (pág. 195) para poder ser clonado.

En este caso no es indiferente la elección de las células de partida, pues cada tipo celular expresa un conjunto de proteínas diferente y, por consiguiente, no contiene los mismos mRNAs. Una vez lisadas las células del órgano, tejido,

tipo celular, estadio de desarrollo, etc., elegidos, se purifica el mRNA, generalmente aprovechando la presencia en su molécula de la *cola de poli(A)* (pág. 280). Se emplea para ello cromatografía de afinidad en columna o microesferas magnéticas de afinidad (pág. 136); en ambos casos, la matriz cromatográfica o microesfera magnética llevan unido un oligo(dT). Cuando el mRNA a clonar no es abundante, se puede amplificar mediante PCR, facilitando así su clonado posterior en células.

Para preparar el cDNA se emplea la transcriptasa inversa, de forma análoga a lo estudiado en la PCR (pág. 196):

Detalle de la preparación de los cDNA



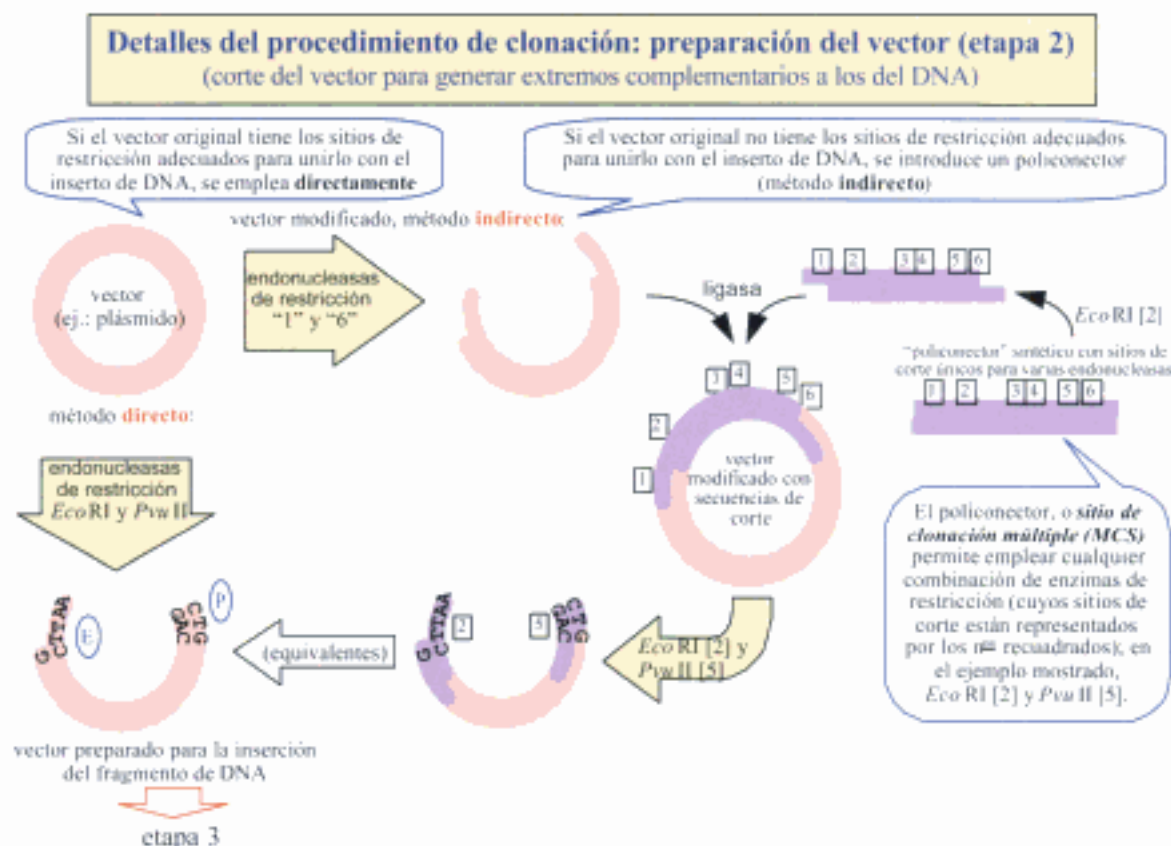
16.2.2 Preparación del vector de clonación

En general, un **vector** es una molécula de DNA de tamaño pequeño, fácil de aislar y caracterizar, con secuencia y mapa de restricción (pág. 226) conocidos, de fácil introducción en la célula anfitriona y una vez allí con capacidad de *replicación autónoma*, es decir, independiente de la replicación del genoma de la célula anfitriona. Es conveniente que el vector posea el mayor número posible de sitios de restricción, para insertar el fragmento de DNA que se quiere clonar, y que incluya al menos un *gen marcador* (por ejemplo, un gen de resistencia a un antibiótico) que permita identificar y/o seleccionar las células que llevan el DNA recombinante (pág. 213). Lo común es que los vectores naturales no posean todas estas características, por lo que habitualmente se emplean vectores modificados (a su vez, obtenidos por técnicas de DNA recombinante, pero ya disponibles comercialmente con todas las prestaciones deseadas).

La misión del vector es unirse con el fragmento de DNA que se quiere clonar (llamado entonces **inserto**) para facilitar su entrada en la célula anfitriona y su replicación. Para poder unir con facilidad ambos, se debe cortar el vector con las mismas enzimas de restricción que se utilizaron para escindir el DNA de la muestra o con enzimas que proporcionen extremos compatibles; de ese modo, sus secuencias complementarias pueden asociarse y ser unidos por la ligasa, dando lugar a una sola molécula de DNA de doble hebra, llamada **DNA recombinante** (rDNA).

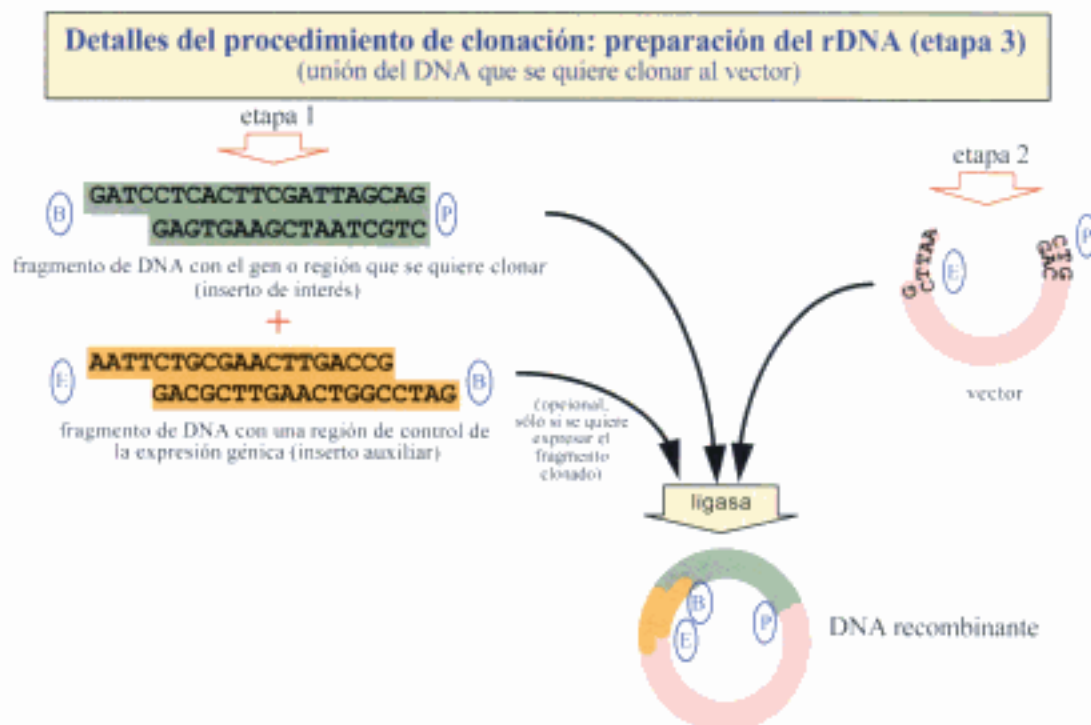
En función de los objetivos de la clonación (amplificación y expresión), se suele hablar de dos grandes grupos de vectores. Por un lado, los **vectores de clonación**, que se emplean únicamente para amplificar el DNA de interés; también se los llama a veces **vectores de inserción** o **vectores de propagación**. Teóricamente, sólo incorporarían el inserto de interés (pág. 205). Por otro lado, los **vectores de expresión**, que poseen además características que facilitan la expresión por la célula anfitriona del DNA de interés. En este caso deberá incorporarse, además de éste, el ya mencionado inserto auxiliar con una región de control de la expresión (pág. 205). En la práctica, los vectores de expresión comerciales suelen incluir ya una región de control, comúnmente un promotor potente que se induzca fácilmente bajo condiciones experimentales controlables, válido para expresar cualquier inserto de interés que se sitúe tras él en el rDNA.

Aunque existen diversos tipos de vectores (pág. 208), por ahora la descripción se hará con un plásmido, una molécula pequeña de DNA circular.

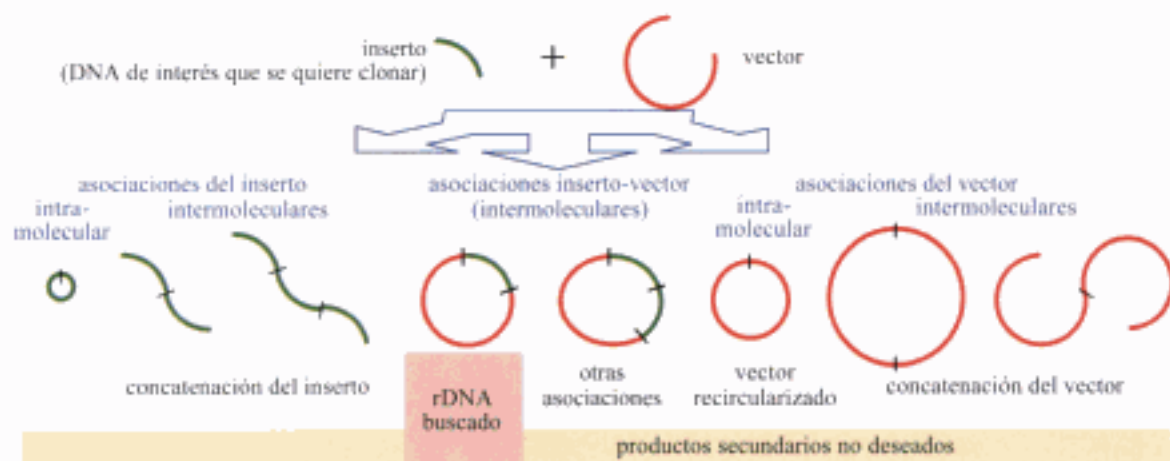


16.2.3 Formación de la molécula de DNA recombinante

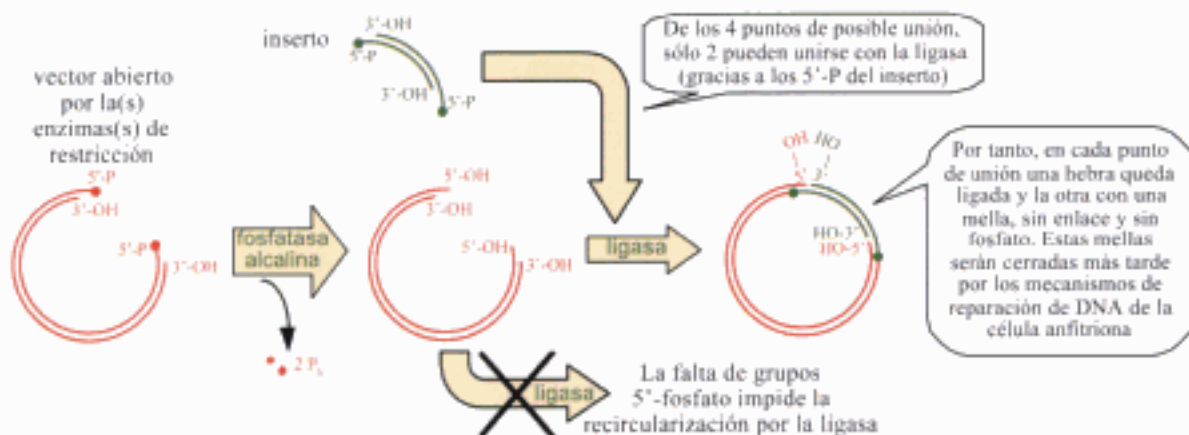
Consiste en la unión covalente, *in vitro*, del inserto de DNA al vector fragmentado, mediante una ligasa. La especificidad de la unión depende del tipo de extremos que hayan resultado al preparar el inserto y el vector: Si son cohesivos, su secuencia es la específica de la enzima de restricción empleada y sólo se asocian los extremos compatibles (pág. 202); una vez asociados por apareamiento de bases, la ligasa establece la unión covalente sellando las dos mellas. Si se trata de extremos romos, no pueden asociarse, pero la ligasa los puede unir, aunque con menor eficacia, y no supone ninguna diferencia la enzima de restricción con la que se prepararon (se puede decir que todos los extremos romos son compatibles).



En la práctica, pueden ocurrir diversos tipos de asociación entre fragmentos de restricción con extremos compatibles, todas ellas catalizadas por la ligasa. En primer lugar, siempre puede darse la asociación **intramolecular**, es decir, entre los dos extremos de una misma molécula (vector o inserto); si se trata de un vector circular, da lugar a su recircularización. Esta interacción está favorecida cuando la concentración de DNA es baja (pues la probabilidad de que dos moléculas se encuentren es menor). En segundo lugar, puede tener lugar una asociación **intermolecular**, tanto la que conduce al rDNA (vector+inserto) como otras que conducen a productos no deseados: vector+vector, inserto+inserto, unión de más de dos moléculas, etc.; todas ellas pueden ser circulares o lineales. Estas interacciones son más probables para concentraciones elevadas de DNA.



Para evitar la formación de estos productos secundarios, que disminuye el rendimiento de la clonación, se acude a estrategias experimentales, previas a la mezcla de vector e inserto, tales como la escisión con dos enzimas de restricción diferentes (pág. 207), que evita la recircularización del vector o, como se describe a continuación, la desfosforilación del vector con fosfatasa alcalina.



16.2.4 Tipos de vectores

Los vectores empleados para la clonación de insertos de DNA son muy variados. Pueden clasificarse según varios criterios, tales como:

- Su procedencia procariótica o eucariótica.
- El tipo de molécula a partir de la que se preparan:
 - Plásmidos: bacterianos, de levadura o de plantas.
 - Virus: que infectan bacterias (bacteriófagos o simplemente fagos), plantas, invertebrados o vertebrados.
 - Cromosomas artificiales: derivados de elementos cromosómicos de fagos (PACs), de bacterias (BACs) o de levaduras (YACs).
 - Quimeras, es decir, moléculas formadas combinando partes de otras cuyo origen es diferente. Normalmente, quimeras de plásmido y fago: cósmidos, fagémidos, fásmidos.
- El tipo de célula anfitriona en la que el DNA recombinante resultante se puede luego incorporar (etapa 4).
- El gen de resistencia que contiene el vector para su posterior detección o selección (etapa 6).
- El tamaño del DNA que admiten como inserto.

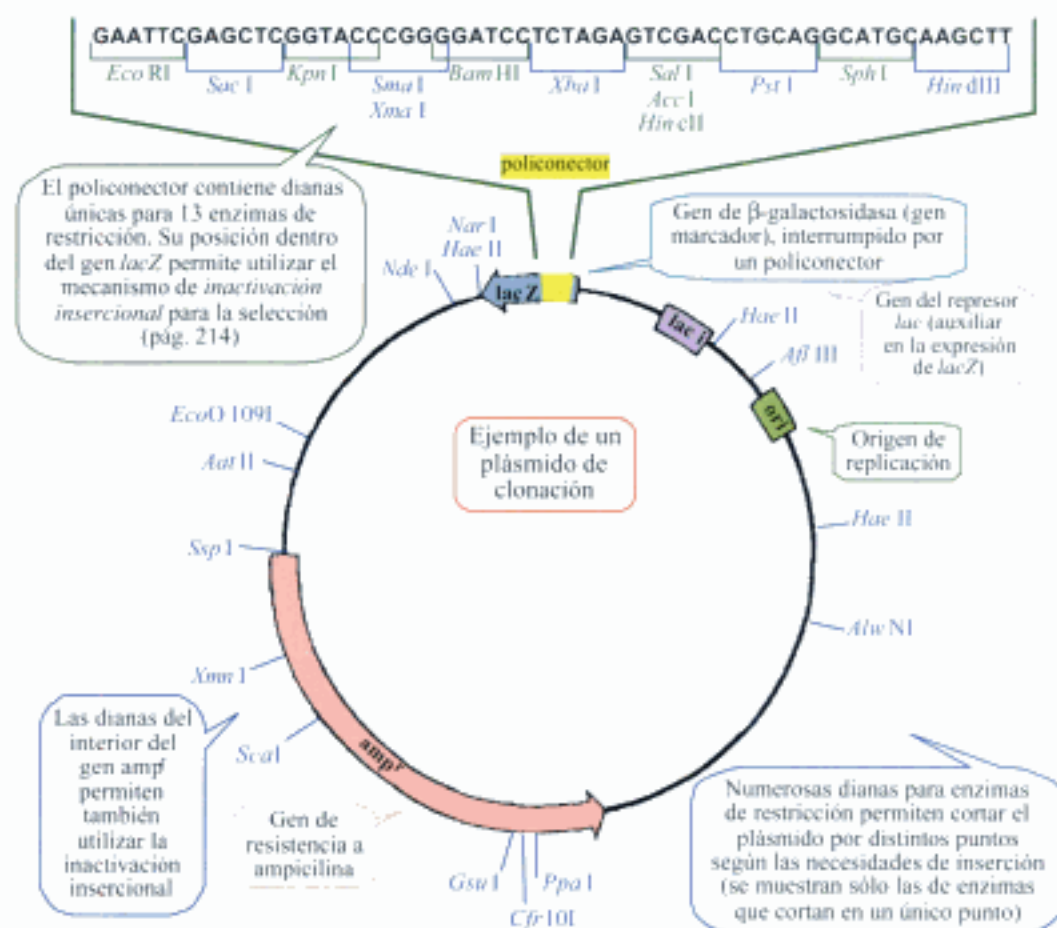
Tipo de vector	Tamaño del inserto que admite	Método de propagación
Plásmidos normales	0-10 kb	Replicación autónoma natural del plásmido
Bacteriófagos, por inserción (fago λ)	0-10 kb	Replicación propia del fago
Bacteriófagos, por sustitución (fago λ)	9-23 kb	Replicación propia del fago
Cósmidos	30-44 kb	Replicación a modo de plásmido
Bacteriófago P1	70-100 kb	
PACs (cromosomas artificiales de P1)	130-150 kb	
BACs (cromosomas artificiales de bacteria)	Hasta 300 kb	
YACs (cromosomas artificiales de levadura)	0,2-1 Mb	

Debe, además, señalarse que continuamente surgen nuevos tipos de vectores como consecuencia del avance en las técnicas de ingeniería genética y las necesidades impuestas por el desarrollo acelerado de los Proyectos Genoma.

16.2.4.1 Plásmidos

Son moléculas de DNA de doble hebra, de pequeño tamaño (2-5 kb) y forma circular, presentes en forma libre en el citosol de muchas bacterias (de 1 a 3.000 plásmidos/célula) y de algunos eucariotas unicelulares (por ejemplo, el plásmido μ de levaduras o "círculo de $2\mu\text{m}$ "). Son fáciles de purificar a partir del cultivo bacteriano. Permiten la incorporación de insertos de DNA de un tamaño máximo de unas 10 kb. Los plásmidos naturales contienen pocos genes propios, ninguno esencial para la viabilidad de la célula, por lo que pueden ser modificados sin afectarla. Para su replicación autónoma (independiente de la del cromosoma bacteriano, aunque utilizando las mismas enzimas), se requiere una secuencia origen de replicación. En cada división celular, el cromosoma origina una sola copia por célula hija, mientras que el plásmido (en su caso, el rDNA formado a partir de él, etapa 3) sufre replications múltiples y origina varias copias por célula (de 20 a 50). Los plásmidos son así replicones (unidades de replicación, pág. 150) que se transmiten y mantienen de manera estable como círculos extracromosómicos.

El empleo de un plásmido como vector de clonación se favorece cuando incluye genes marcadores, cuya presencia pueda ser detectada o cuya expresión confiera a la célula anfitriona portadora del rDNA una serie de propiedades especiales (pág. 214). Por ello, se utilizan comúnmente **plásmidos artificiales** o sintéticos, obtenidos a su vez por técnicas de DNA recombinante y disponibles comercialmente. Baste como ejemplo el plásmido **pUC19**.

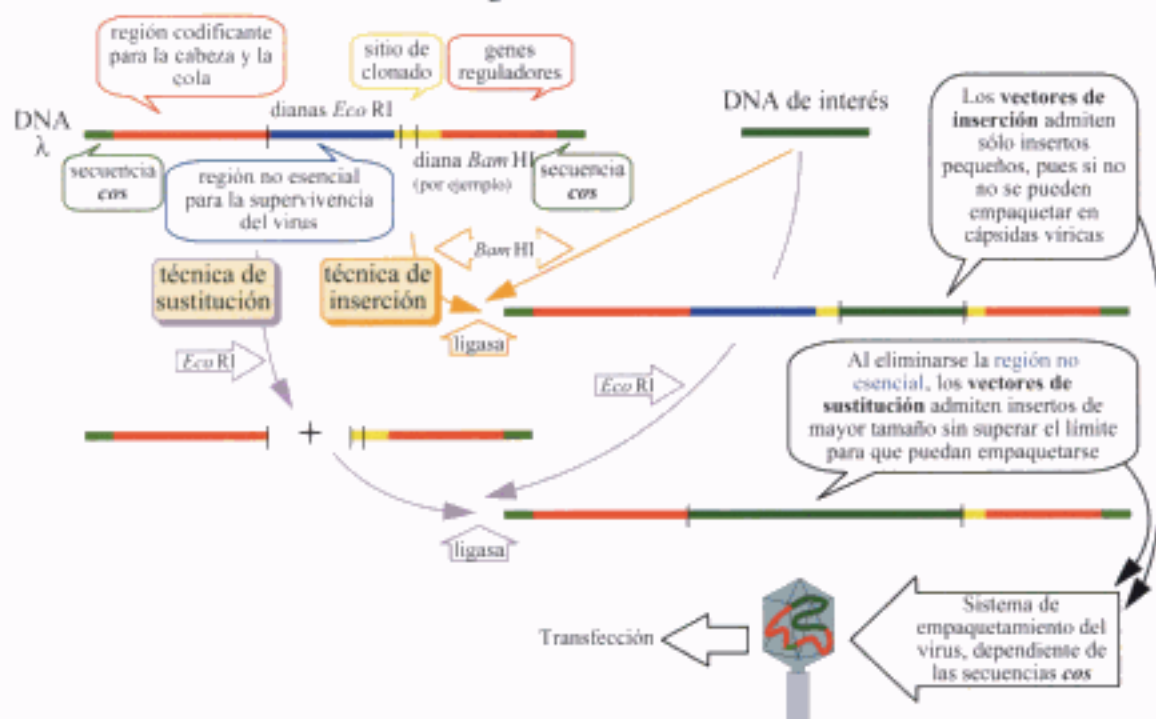


16.2.4.2 Bacteriófagos

Los virus en general se replican mediante la introducción de su genoma en células procariotas o eucariotas (infección). Esta propiedad resulta útil para conseguir una alta eficiencia de clonación, empleando como vectores virus modificados. Cada tipo de virus infecta a células específicas y, dependiendo del tamaño del genoma vírico, admite insertos más o menos grandes, pero en general mayores que los que se pueden insertar en plásmidos. Así, para clonar en bacterias se usan **bacteriófagos**, **baculovirus** para células de insecto, y virus **SV40** o **retrovirus** para células de mamífero. La proliferación de los virus recombinantes se consigue en un cultivo de las células anfitrionas respectivas, en las que el virus se reproduce. El uso de virus es especialmente adecuado para la clonación en células de mamífero, por su elevada eficacia de introducción del rDNA, pero dada su complejidad y por razones de extensión no se tratará en este tema.

Los virus más comúnmente empleados como vectores son los parásitos bacterianos **fagos M13** y λ (**lambda**). El tamaño del vector recombinante se ve limitado por la necesidad de su empaquetamiento dentro de la cápsida. El inserto se puede introducir de la misma forma que para los vectores plasmidicos, es decir, cortando inserto y DNA del fago con las mismas enzimas de restricción y ligando para obtener un rDNA; esto se denomina *técnica de inserción*, y existen vectores específicos para ello. Cuando el tamaño del DNA de interés que se quiere clonar es mayor, se puede incorporar eliminando previamente la parte del genoma vírico que no se requiere para la infección y replicación; se habla entonces de *técnica de sustitución o reemplazamiento*.

Vectores de clonación basados en el fago lambda



16.2.4.3 Cósmidos

Son **vectores sintéticos**, quimeras que combinan características de plásmidos y fagos para combinar las ventajas de ambos tipos de vectores. El **fragmento plasmidico** proporciona las características de tamaño pequeño (5-7 kb), circularidad, un origen de replicación, un número de sitios de restricción donde se va a insertar el DNA, uno o más marcadores seleccionables y la propagación en bacterias como si fuesen plásmidos. Del **fago λ** se toman las secuencias "cos" (del inglés *cohesive sites*), que proporcionan al cósmido la capacidad, propia del DNA del fago, de ser empaquetado en cápsidas víricas *in vitro*.

Los cósmidos son útiles para clonar insertos de gran tamaño (hasta 45 kb), especialmente en la preparación de librerías genómicas de organismos superiores, ya que disminuyen en gran medida el número de clones necesarios para que todo el genoma esté representado en la genoteca (pág. 216). Los métodos de introducción del rDNA cósmidico (etapa 4) dependen del tamaño del inserto. Si éste es pequeño, se puede introducir por transformación (las células se hacen competentes para incorporarlo en forma de plásmido recombinante, pág. 211). Si el tamaño es relativamente grande, la transformación se complica y debe acudir al empaquetamiento previo en forma de fagos, aprovechando las secuencias cos.

16.2.4.4 Cromosomas artificiales

Este tipo de vectores se han diseñado en especial para adaptarse al gran tamaño de insertos requeridos en el caso del genoma humano y de otros mamíferos, y para la clonación en células eucariotas.

a) Cromosomas artificiales bacterianos

Se les llama comúnmente BACs (*bacterial artificial chromosomes*). Se trata de vectores sintéticos derivados del plásmido F o el fago P1. Se caracterizan, en primer lugar, por aceptar grandes insertos de DNA, entre 100 y 300 kb. En comparación con otros tipos de vectores, los BACs contienen genes que reducen su capacidad de copia, es decir, de replicación en la célula anfitriona del rDNA formado (etapa 4). Como resultado, se consigue mayor estabilidad de los insertos de DNA de origen eucariótico, que en cósmidos sufren con frecuencia reordenamientos internos. Gracias a estas dos características, los BACs son vectores adecuados para la elaboración de genotecas eucarióticas (pág. 215).

b) Cromosomas artificiales de levadura

Los YACs (*yeast artificial chromosomes*) son vectores sintéticos que contienen las tres regiones relacionadas con la funcionalidad de un cromosoma (pág. 92): secuencias constituyentes de un centrómero, de dos telómeros y de un origen de replicación. Además, incorporan en general genes marcadores seleccionables (etapa 6) y un sitio de clonado múltiple (etapa 2) en el que admiten un inserto de gran tamaño, que puede alcanzar 1 Mb, lo que constituye una característica fundamental. En ocasiones, este tipo de vectores se denominan ARS/CEN/TEL, por los 3 tipos de secuencias que contienen. La secuencia centromérica permite la segregación correcta de las copias hijas del cromosoma durante la división de la célula (pág. 92), mientras que los telómeros estabilizan, como ya se ha estudiado, los cromosomas a lo largo de sucesivas divisiones (pág. 157). El origen de replicación o secuencia de replicación autónoma (ARS, *autonomously replicating sequence*) permite la formación de copias del cromosoma artificial de forma independiente a la del genoma propio de la célula anfitriona. Los YACs se utilizan en particular para preparar genotecas de genomas muy grandes, o para mantener un gen completo en un único clon.

16.2.5 Incorporación del DNA recombinante a la célula anfitriona

La replicación de las moléculas de rDNA (clonación) requiere que éstas se introduzcan en una **célula anfitriona** (receptora u hospedadora, incorrectamente llamada *célula huésped*). Para que esta incorporación, que en general tiene lugar por mecanismos poco conocidos, sea eficaz es preciso emplear células especializadas (con un genotipo seleccionado), adecuadas al vector de clonación utilizado y con capacidad de división.

16.2.5.1 Tipos de célula anfitriona

Con frecuencia, las células anfitrionas son **procarióticas**, principalmente cepas de *E. coli*, por ser fáciles de obtener, manipular, multiplicar, y conocerse su genética. Al proceso de introducción del rDNA plasmídico se le denomina **transformación** (pág. 212). La(s) molécula(s) de rDNA plasmídico incorporadas se replican luego varias veces dentro de la célula, de acuerdo con las propiedades de los plásmidos. El inserto de DNA puede permanecer unido al plásmido (lo más frecuente) o recombinarse con el DNA del cromosoma. Si se integra en el vector más de un tipo de DNA, por ejemplo, dos genes o un gen y un regulador de la transcripción, el proceso de incorporación a la célula anfitriona del rDNA se llama **cotransformación**.

Las **células anfitrionas eucarióticas** son más difíciles de obtener y mantener, por lo que sólo se utilizan con fines específicos, fundamentalmente para el análisis de la regulación génica, la expresión de una proteína eucariótica, etc. De entre ellas, las más manejables en cultivo son las **levaduras** que, aunque mantienen algunas características similares a procariotas (células individuales, crecimiento rápido y económico, manipulables, etc.), tienen una organización celular propia de eucariotas y son de interés biotecnológico por su capacidad de expresión y maduración de proteínas.

Entre las **células animales** que pueden establecerse en cultivo, se pueden citar las células HeLa de carcinoma epidérmico humano, los fibroblastos NIH de ratón y BHK de riñón de hámster, los ovocitos de ratón, etc. En general, pueden dividirse multitud de veces conservando los mismos caracteres, por lo que su rendimiento de clonación es elevado.

16.2.5.2 Métodos de introducción del rDNA

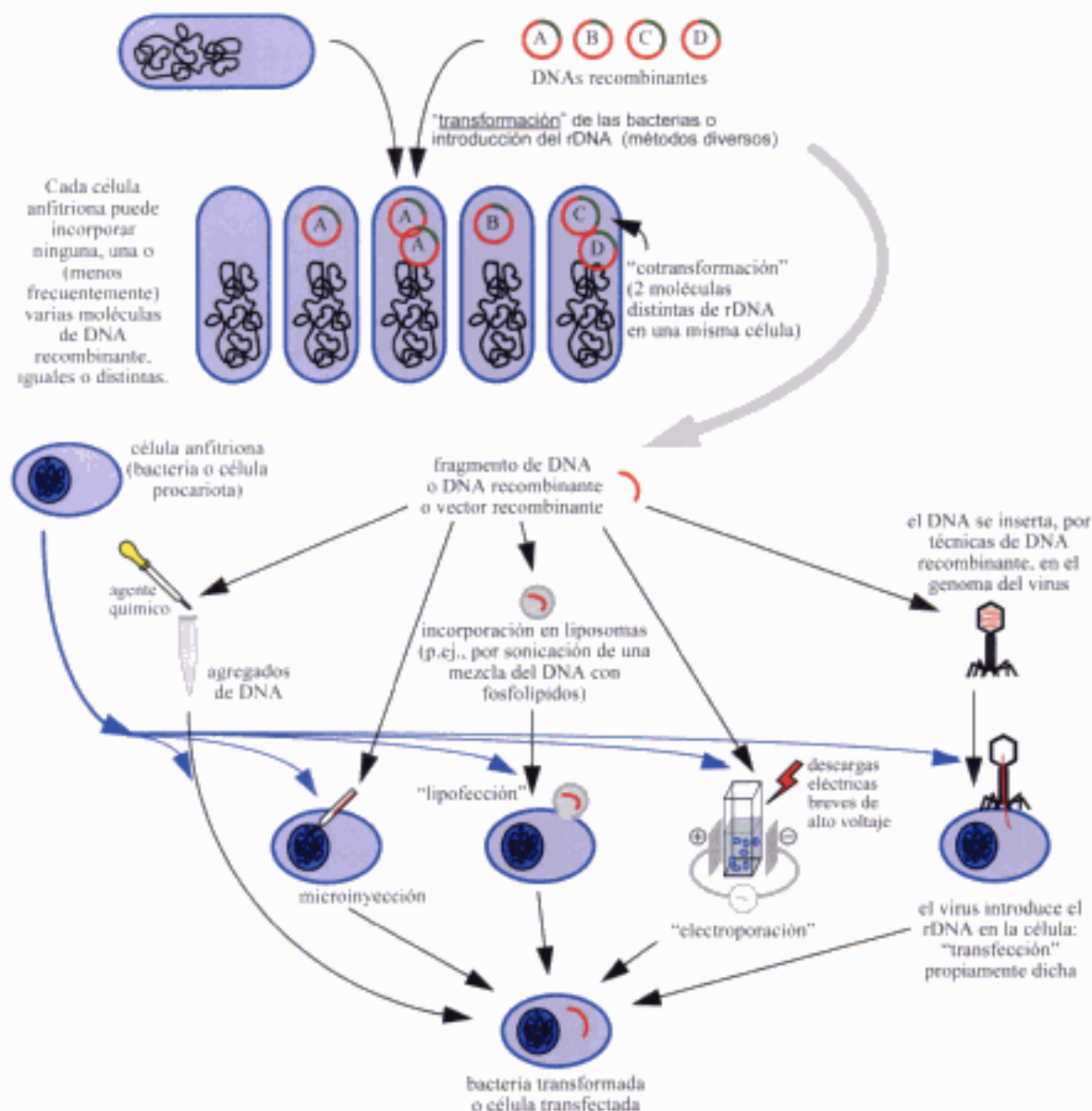
Puesto que normalmente la membrana plasmática muestra una permeabilidad selectiva y no admite grandes fragmentos de rDNA, se acude a diversos métodos para facilitar su captación por alteración transitoria de la permeabilidad celular (métodos pasivos) o introducir el rDNA directamente (métodos activos). En muchos casos, los mecanismos son poco conocidos y los métodos resultan más o menos eficaces dependiendo del tipo de vector y de célula. El proceso suele denominarse **transformación** o **transfección**, aunque estrictamente el primer término corresponde a la introducción en bacterias anfitrionas y el segundo al empleo de virus como vectores. Los métodos utilizados pueden dividirse también en físicos y químicos.

- El tratamiento más frecuente es de tipo físico-químico y se dice que la célula se hace **competente**. Se comienza con una incubación en presencia de CaCl_2 a 0 °C, seguida de un cambio brusco a 37-43 °C, que induce a las

bacterias a aceptar moléculas de DNA. Se dispone de cepas establecidas especialmente susceptibles a este proceso.

- La incorporación pasiva de DNA se favorece cuando sus moléculas se han agregado o coprecipitado por interacción con **catiónes**, típicamente fosfato cálcico o DEAE-dextrano (dietilaminoetil-dextrano, un polímero policationico).
- Otros agentes químicos, en especial el polietilenglicol (PEG), aumentan la permeabilidad de la membrana celular, con lo que se incorpora más fácilmente el rDNA, también de forma pasiva.
- El DNA se puede incorporar en **liposomas**, que luego con facilidad se fusionan con la membrana celular y liberan su contenido al interior (incluyendo el DNA que, en realidad, parece asociarse a la superficie exterior hidrofílica del liposoma).
- Un método con gran eficacia es el empleo de vectores de tipo **vírico**, que por su vía natural de infección introducen en la célula anfitriona su DNA, en este caso recombinante. Para este tipo de incorporación se ha acuñado el término *transfección*.

Detalles del procedimiento de clonado: incorporación del rDNA (etapa 4)



- Aplicando **descargas eléctricas** en forma de pulsos breves de alto voltaje se consigue la apertura de poros transitorios en la membrana, por los que puede pasar el rDNA. Este método de *electroporación* es bastante habitual para células eucarióticas, aunque de resultados muy variables.
- Un método curioso y de aplicación creciente, la **"biolística"**, consiste en el **bombardero** de las células anfitrionas con microesferas (de tungsteno u oro) recubiertas con el rDNA. Se emplean dispositivos adecuados ("pistolas de genes") que impulsan esos microproyectiles mediante un gas a presión, pudiendo aplicarse sobre suspensiones

celulares, células en cultivo o incluso *in vivo* (sobre la piel, hojas ...). Algunos de los proyectiles impactan y penetran en las células, cuya membrana se resella rápidamente, quedando el DNA incorporado en su interior.

- En casos donde la eficiencia de suministro del rDNA debe ser máxima se acude a la **inyección directa** del DNA en la célula o en su núcleo, empleando micromanipuladores y agujas de vidrio (microjeringas o micropipetas, de punta extremadamente fina), bajo observación al microscopio. Obviamente, el inconveniente es el escaso número de células que se pueden tratar por este método manual, por lo que generalmente se utiliza sólo sobre oocitos y cigotos. Dado su carácter directo, se puede incluso prescindir del vector y suministrar el fragmento de DNA deseado sin modificar.

Agente promotor de la incorporación	Nombre específico del método	Carácter	
Ca ²⁺ (generalmente como fosfato)	transformación, células competentes	químico	pasivo
DEAE-dextrano		químico	pasivo
PEG (polietilenglicol)		químico	pasivo
liposomas	lipofección	físico	pasivo
vectores virales	transfección	físico	activo
descargas eléctricas	electroporación	físico	activo
microproyectiles	biolística	físico	activo
inyección directa con micropipetas	microinyección	físico	activo

16.2.6 Propagación en cultivo

La **división** de las células anfitrionas en cultivo producirá, junto con la replicación de su propio genoma, la formación de copias del rDNA que haya podido incorporar y, por tanto, la **clonación** propiamente dicha. En la práctica, se suele hacer una primera siembra celular sobre una placa de agar para que crezcan las colonias (cada colonia proviene de una célula, son clones) y una resiembra posterior de las colonias en medio líquido, que forma poblaciones con múltiples copias de cada célula original (clones celulares). Dependiendo de las propiedades del vector, el rDNA puede replicarse de forma rápida e independiente del genoma de la célula anfitriona.

Detalles del procedimiento de clonado: propagación celular (etapa 5)



Simultáneamente con esta etapa de propagación suele tener lugar la selección de clones recombinantes y la expresión opcional del gen clonado.

16.2.7 Detección y selección de los clones recombinantes

Dado el gran número de células empleadas en la clonación, y la diversidad de productos que pueden originarse en el proceso de creación del DNA recombinante, es esencial poder diferenciar qué células (o colonias a las que han dado lugar en cultivo) han incorporado realmente el DNA recombinante deseado. Esta **detección** se puede combinar con una **selección**, es decir, conseguir que sólo las células adecuadas sobrevivan en el cultivo. Esto facilita mucho la tarea de propagación celular, reduciendo el número de clones que se deben seguir cultivando. Dependiendo del método utilizado, es posible distinguir células que han incorporado el vector (recombinante o no), que han incorporado rDNA o incluso que han incorporado el rDNA con el inserto dispuesto correctamente.

Se pueden considerar tres tipos de métodos de detección:

- **Métodos de hibridación:** detección de una secuencia del DNA clonado, o del mRNA transcrito a partir de él, por hibridación con una sonda marcada.
- **Métodos inmunológicos:** detección de la proteína codificada por el gen clonado, mediante un anticuerpo específico.

• **Métodos genéticos o de selección fenotípica:** en ciertos casos, la expresión del propio inserto produce cambios apreciables en la célula anfitriona, que sirven para reflejar su clonación con éxito. Sin embargo, la estrategia más común es la detección y selección basadas en la presencia de **genes marcadores** en el vector:

- **Genes que codifican una proteína detectable**, comúnmente una enzima para la que se dispone de un ensayo. Por ejemplo, la β -galactosidasa (gen *lacZ* de *E. coli*) transforma el sustrato sintético X-gal (5-bromo-4-cloro-3-indolil- β -D-galactopiranosido) en un producto de color azul.
- **Genes de resistencia a un antibiótico**, que codifican una proteína que permite el crecimiento de la célula anfitriona en presencia de un determinado antibiótico. Por ejemplo, el gen *amp^r* codifica una β -lactamasa capaz de degradar la ampicilina, una penicilina semisintética. El gen de resistencia a tetraciclina (*tetA*) codifica una proteína de membrana que actúa como una bomba, extrayendo este antibiótico de la célula.
- **Genes que complementan defectos nutricionales**. En este caso, se emplean cepas mutantes de la célula anfitriona, que no pueden crecer en ausencia de un nutriente. Por ejemplo, clonando en levaduras con un defecto metabólico en la ruta de biosíntesis de triptófano, se emplea un vector que contiene el gen *TRP1* correcto, con lo que sólo las células transformadas crecen sin Trp en el medio.

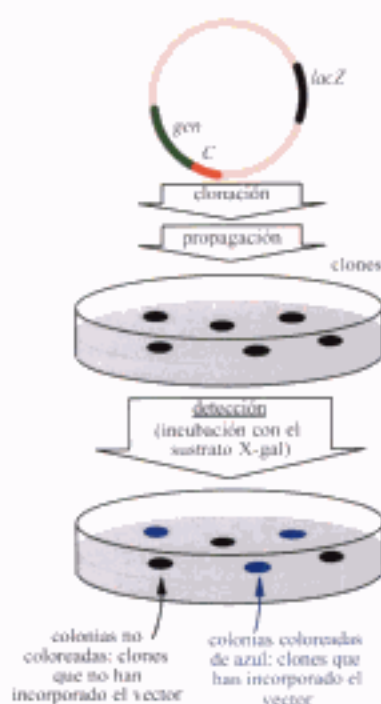
Los planteamientos experimentales son muy diversos y generalmente combinan varios de estos marcadores. Por razones didácticas, se recogen unos ejemplos que ilustran el uso de varios genes marcadores, aunque no se ajusten a la práctica del laboratorio. Un concepto importante a este respecto es el de **inactivación insercional**: la inclusión de un policonector o sitio de clonado múltiple (págs. 207, 209) en el interior de un gen marcador hace que la acción de éste se ejerza sólo en las células que han incorporado vector sin inserto.

Detalles del procedimiento de clonación: detección y selección de clones recombinantes (etapa 6)

Ejemplo de un vector para clonación: contiene el inserto de DNA (gen de interés), una región de control y un gen auxiliar (gen marcador o gen de resistencia a antibióticos)

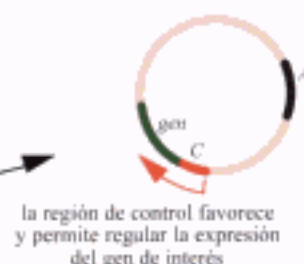


Detección de clones:
el gen A sería un gen marcador, codifica una proteína que se puede detectar

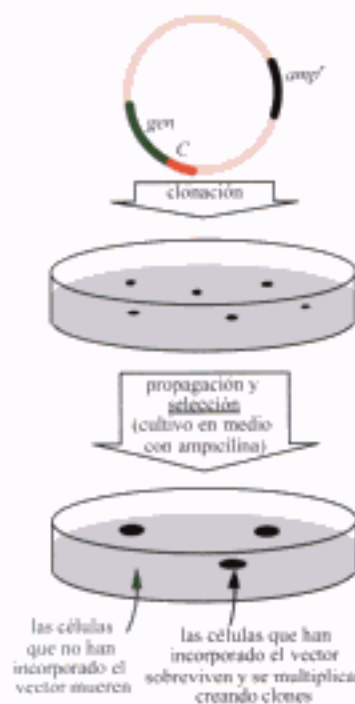


fragmento de DNA con el gen que se quiere clonar
fragmento de DNA con una región de control (promotores de transcripción)

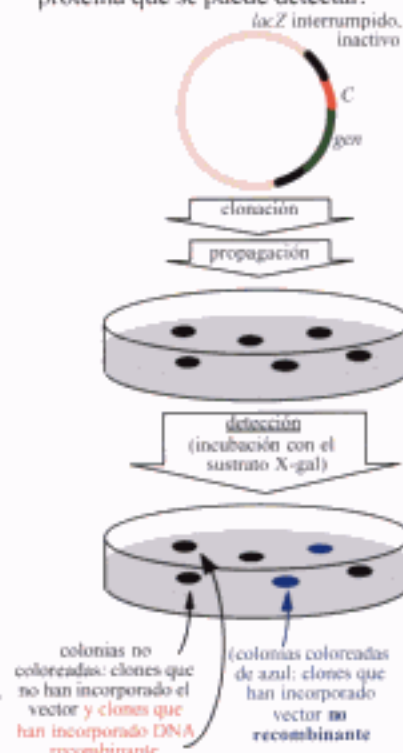
restricción y ligamiento



Selección de clones:
(también sirve como detección)
el gen A sería un gen de resistencia a antibiótico



Detección de clones por inactivación insercional: el gen A sería un gen marcador, codifica una proteína que se puede detectar.



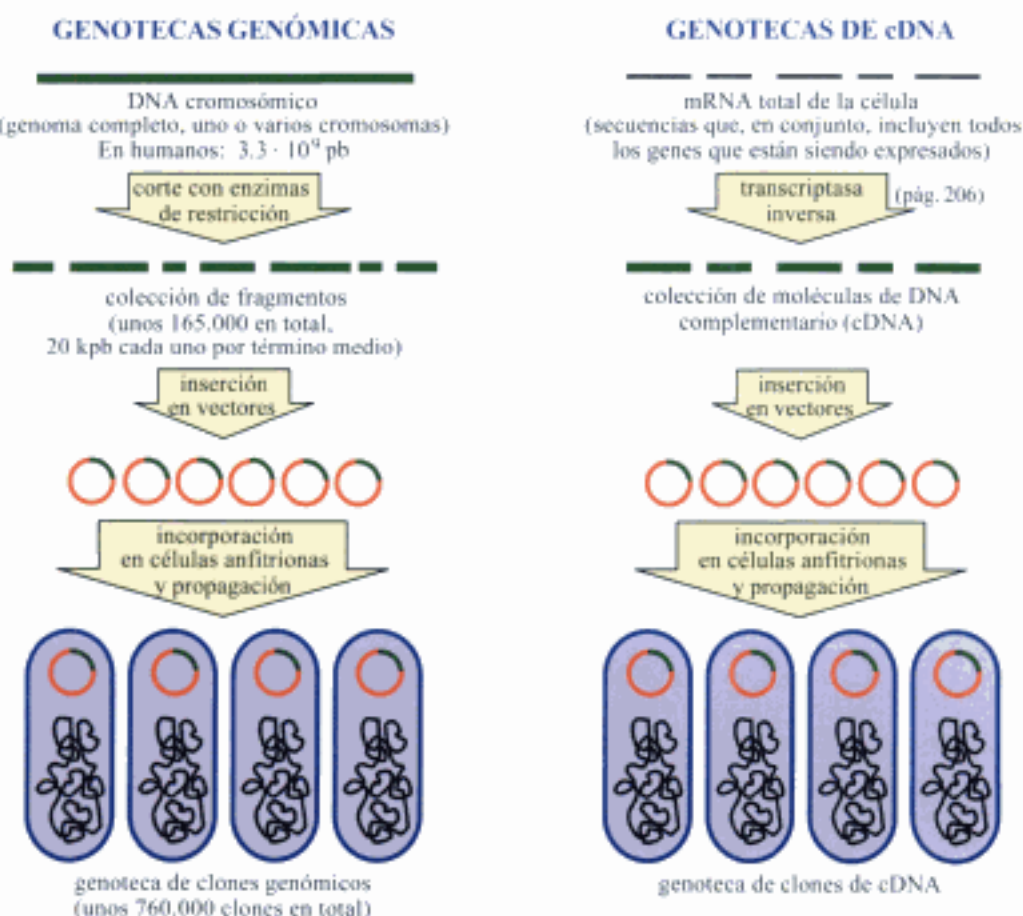
16.2.8 Expresión génica

En algunos casos se pretende no sólo la amplificación del DNA clonado, sino su expresión, bien para estudiar este proceso o para investigar el producto (RNA o proteína). Lógicamente, el DNA clonado será en estos casos un gen, y por ello con frecuencia se tratará de un cDNA. Para conseguir esta expresión se emplean **vectores de expresión**, especializados, distintos de los vectores de clonación por presentar características que favorecen la transcripción del gen una vez incorporado a la célula anfitriona. Típicamente, un vector de expresión incluye (o se debe insertar en él a la vez que el gen, pág. 205) un **promotor**, región reguladora de la transcripción (pág. 264). Es frecuente emplear promotores inducibles, ajenos al organismo del que procede el gen clonado pero que se activan bajo condiciones experimentales controlables, como la adición de un nutriente u otro tipo de compuesto, disparando entonces la transcripción del gen clonado.

16.3 GENOTECAS

Se llama genéricamente **genoteca** o **biblioteca de DNA** a una colección de clones formados a partir de todos los fragmentos de DNA obtenidos previamente por escisión del genoma de una especie o de un tejido. En el caso de la clonación celular, cada uno de estos fragmentos se presenta integrado, bajo la forma de rDNA, en un vector incorporado en uno de los clones celulares que componen la genoteca. En la clonación acelular se presentan como fragmentos individuales independientes.

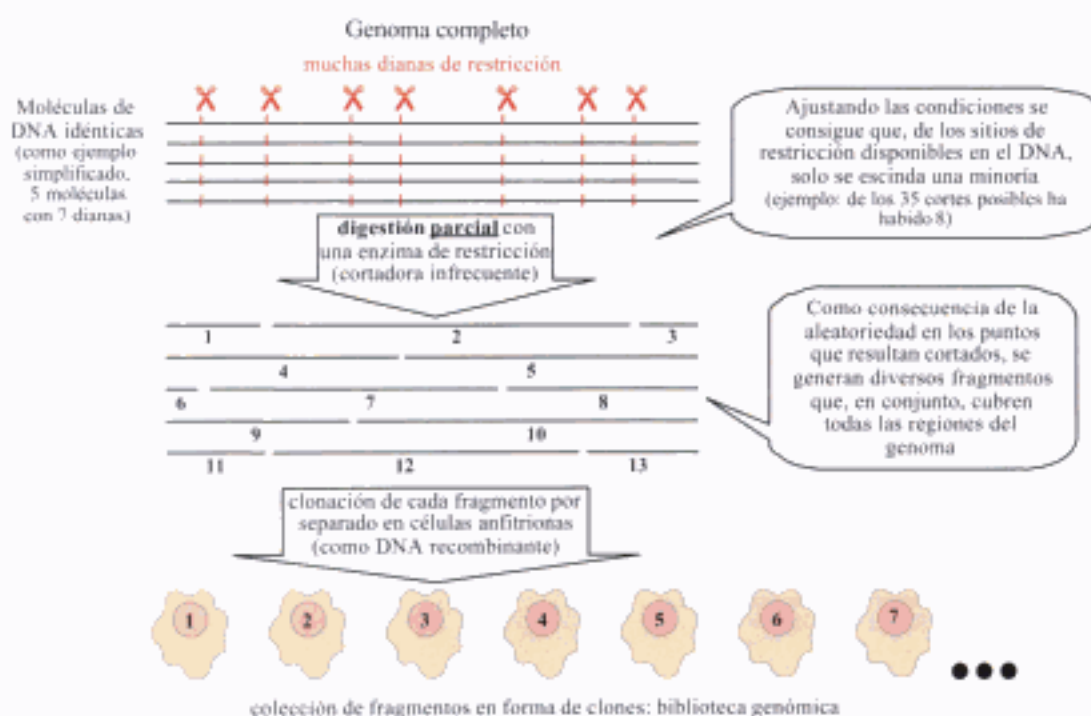
Dependiendo de si las muestras de partida para la clonación son fragmentos del DNA genómico o moléculas de mRNA (pág. 205), se distingue entre **genotecas genómicas** y **genotecas de cDNA**. Es importante indicar que la información proporcionada por ambas genotecas no es idéntica, pues sólo las primeras incluyen tanto las regiones codificantes como las no codificantes; además, las genotecas de cDNA son diferentes según el tejido particular del que se han preparado, pues sólo contienen las regiones codificantes que se expresan en ese tejido, tipo celular, etapa del desarrollo o diferenciación, entorno bioquímico... De acuerdo con ello, se utilizan alternativamente distintas genotecas como punto de partida para la clonación, secuenciación, etc. del DNA.



Una vez preparada una genoteca, ésta se convierte en el punto de partida para la búsqueda de cualquier región del DNA o de genes y su estudio mediante clonación adicional (que recibe el nombre de *subclonación*), secuenciación, expresión, etc. Con respecto al uso de las genotecas como fuente para otros estudios, el problema genérico que se plantea es la localización en la genoteca del clon que contiene el DNA buscado. Para ello se acude a los distintos métodos de detección y selección, previamente comentados al describir la clonación de un solo fragmento (pág. 213).

16.3.1 Genotecas genómicas

El concepto clásico de genoteca corresponde en realidad estrictamente al de genoteca genómica, representativa del genoma completo. El número de clones componentes debe ser, en la práctica, suficiente para abarcar todo el genoma, incluyendo tanto el DNA codificante como el no codificante. La preparación del DNA de partida se realiza por **escisión** del genoma con enzimas de restricción bajo condiciones de rotura **parcial**, es decir, que no permitan la actuación exhaustiva de la enzima sobre todos los sitios de restricción presentes en el DNA. La rotura se produce en cada molécula aleatoriamente sobre el conjunto de sitios de restricción, de forma que aparecen un número menor de fragmentos, y éstos son de mayor tamaño, respecto a lo teóricamente posible, y diferentes de una a otra molécula.



A partir del tamaño del genoma completo y el tamaño medio de los insertos se puede calcular el mínimo número de clones necesario para que la genoteca represente al genoma completo. Sin embargo, el carácter aleatorio de la preparación de fragmentos y del proceso de clonación hacen que, para tener certeza de que una genoteca contiene cualquier región del genoma, se requiera en la práctica un número de clones muy superior al teórico. Este número de clones reales dividido por el número de clones teóricos define el parámetro conocido como **equivalente genómico (GE)**, que es de 3 a 10 para conseguir un 99% de probabilidad de clonación.

Relación entre tamaños de genoma e inserto y número de clones teóricos y prácticos

Organismo	Tamaño del genoma (Mb)	Tamaño de inserto (kb)	Nº de clones independientes (teóricos)	Nº de clones en la práctica	Equivalente genómico (EG)
Bacterias	~ 4	20	$4.000 / 20 = 210$	~ 900	$900 / 210 = 4$
Levaduras	~ 14	20	$14.000 / 20 = 700$	~ 3.500	$3.500 / 700 = 5$
Humanos	~ 3.300	20	$3.300.000 / 20 = 165.000$	~ 760.000	$760.000 / 165.000 = 4,6$

16.3.2 Genotecas de DNA complementario

A diferencia del caso anterior, las genotecas de cDNA representan exclusivamente a la parte codificante del genoma. Son especialmente útiles en eucariotas, donde una célula sólo expresa una pequeña fracción de su genoma (menos del 3% en humanos), lo que reduce el tamaño del material genético disponible para construir la genoteca y, por tanto, el número de clones. El material de partida, el mRNA de la célula, contiene menos de 20.000 moléculas diferentes, con un tamaño medio de 2 a 3 kb. Por otro lado, puesto que el número de copias de los distintos tipos de mRNA es muy variable en eucariotas, su representación en la genoteca no será uniforme. Puesto que los cDNAs carecen de señales de expresión, cuando se pretenda la expresión del gen clonado es imprescindible incluir en el vector de expresión un potente promotor adecuado para la célula anfitriona, de modo que ésta exprese la proteína codificada por el cDNA.

Tema 17

Genómica. Obtención de los mapas genético y físico del genoma

17.1	Introducción	217
17.2	Mapas genéticos	218
17.2.1	Fracción de recombinación como medida de distancia en un mapa genético	218
17.2.2	Obtención del mapa genético	220
17.3	Mapas físicos	221
17.3.1	Obtención de mapas físicos a baja resolución	221
17.3.1.1	Métodos basados en empleo de líneas celulares somáticas híbridas	221
17.3.1.2	Obtención de mapas mediante ensayos de hibridación con sondas fluorescentes	224
a)	Mapas por hibridación <i>in situ</i> (FISH) cromosómica	224
b)	Mapas por citometría de flujo	225
17.3.2	Mapas físicos de alta resolución	225
17.3.2.1	Mapa físico por hibridación <i>in situ</i> (FISH) de alta resolución	225
17.3.2.2	Mapa físico con enzimas de restricción	226
17.3.2.3	Mapas físicos de STS y EST	228
17.3.2.4	Mapas de contigios	229

17.1 INTRODUCCIÓN

Se conoce hoy día como **genómica** al conjunto de estrategias y tecnologías empleadas para la caracterización molecular del genoma, en su más amplio sentido, pero con un énfasis especial en el genoma humano. Conceptualmente, la genómica corresponde a la obtención del **mapa de un genoma**, que a su vez puede ser considerado como la **estructura fina** del material genético de cualquier especie. Por tanto, constituye la base de una Biología Molecular e Ingeniería Genética modernas aplicadas a la Sanidad.

De acuerdo con estos objetivos, se pueden distinguir dos **áreas** dentro de la Genómica, consecutivas, aparentemente independientes pero totalmente relacionadas entre sí:

En primer lugar, la **genómica estructural**, que corresponde al estudio de las técnicas y estrategias necesarias para obtener tanto los mapas físicos del genoma, posibles a distintos niveles de resolución, como la secuencia total del DNA genómico, o mapa físico a la resolución máxima, el nucleótido individual. Los mapas físicos, tanto clásicos como modernos, se describen en el presente tema, mientras que la secuenciación propiamente dicha, o descifrado del conjunto de "letras" que constituyen la información básica de cualquier organismo (3.300 millones de pb en humanos), se describen en el tema 18. La mayoría de los proyectos genoma en curso se encuentran dentro de alguno de estos estadios de la genómica estructural. En el caso del **Proyecto Genoma Humano (PGH)**, se está llegando a lo largo del año 2000 al último eslabón. De ahí la enorme trascendencia de todos los proyectos implicados en este gran objetivo.

En segundo lugar, la **genómica funcional**, correspondiente a la caracterización del **proteoma**, es decir, el conjunto de todas las proteínas originadas por el genoma, bajo los diversos patrones posibles de expresión génica a proteínas, modificación y degradación proteica y desarrollo de su función, todo ello dentro de cada tipo de célula, momento celular, condiciones particulares *in vivo*, etc. Corresponde, en definitiva, al estudio funcional del DNA codificante. Por analogía, se llama **proteómica** al conjunto de técnicas y estrategias utilizadas para el estudio del proteoma. En contraste con la genómica estructural, la proteómica se encuentra en sus inicios, y sólo se tienen datos para algunos organismos modelo. Transcurrirán años antes de conocer al completo la función de todos los genes humanos y sobre todo para dar utilidad práctica, en su caso, a las proteínas, tanto con fines básicos como aplicados en favor de la humanidad (por ejemplo a la biomedicina).

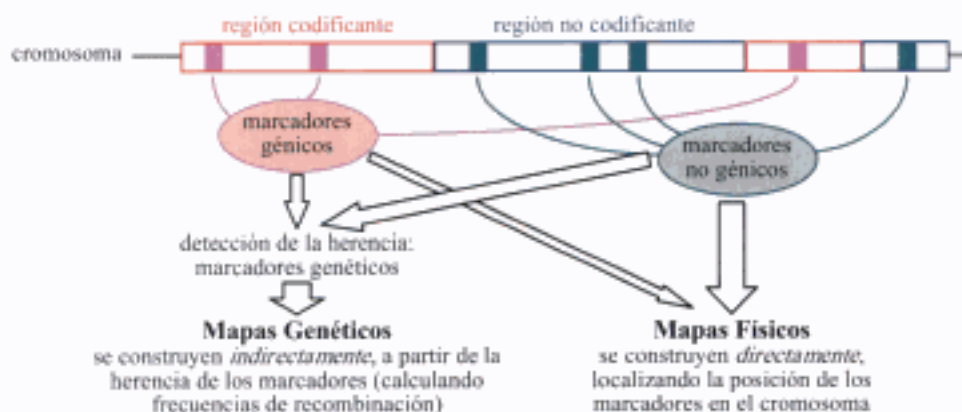
El PGH es un proyecto de tal envergadura que sus objetivos no se pueden alcanzar con el simple planteamiento de buscar métodos para la etapa de secuenciación, a pesar de su creciente eficiencia. De hecho, el progreso logrado ha supuesto la implicación de numerosas iniciativas y metodologías que superan con creces dicha etapa. Ante la imposibilidad material de recoger toda esta información (por otro lado, ya disponible de forma especializada día a día en Internet), se presentan en este tema y el siguiente las líneas básicas para un enfoque "lógico" del conocimiento del genoma.

La elaboración del mapa, o "mapeo", de un genoma, tanto el humano como los de otros organismos, se ha abordado a lo largo del tiempo en varias etapas consecutivas, a niveles crecientes de resolución. Éstos se corresponden, a grandes rasgos, con tres grandes estrategias metodológicas, complementarias entre sí: **mapeo genético**, **mapeo físico** y **secuenciación**. Además ha sido necesario el desarrollo de técnicas informáticas para recoger, almacenar, distribuir y analizar el creciente volumen de datos obtenidos.

Fases del PGH		Características
Mapeo	Mapa genético o de ligamiento	Proporciona la menor resolución. Sitúa de forma aproximada en los cromosomas diversos loci (de genes y de otros marcadores o secuencias identificables). El cálculo de la distancia y, por tanto, la ubicación relativa de genes o marcadores, se basa en la frecuencia con la que los pares de loci se transmiten conjuntamente a la descendencia.
	Mapa físico	Asigna localizaciones a los marcadores, definidas por distancias en pares de bases. Destacan entre sus variantes los mapas de restricción (ordenación en el cromosoma de las dianas para enzimas de restricción).
Secuenciación		Equivale a la obtención del mapa genómico en su nivel más detallado. Su objetivo es establecer la secuencia nucleotídica completa del genoma humano, explorando los 3.300 millones de pb de la célula haploide (6.600 en la célula diploide). Se lleva a cabo en paralelo y con el apoyo de la información proporcionada por los mapeos genético y físico. Hace uso de planteamientos y métodos sumamente específicos, cada vez más automatizados e informatizados. Se describe en el tema 18.

17.2 MAPAS GENÉTICOS

Inicialmente, los mapas genéticos consistían en una representación gráfica de la ordenación y ubicación de los genes en cada cromosoma. Sin embargo, actualmente se incluyen en estos mapas no sólo genes, sino otras regiones del DNA, de modo que para englobar a ambos se emplea el término **marcador genético**; este papel puede desempeñarlo cualquier característica física o molecular del DNA que difiera entre individuos y sea detectable en el laboratorio, para poder seguir su patrón de herencia. Como marcadores genéticos, por tanto, se incluyen tanto regiones que forman parte de un gen (**marcadores génicos**) como segmentos no codificantes (**marcadores no génicos**). La gran mayoría de marcadores genéticos en los mapas actuales son marcadores no génicos.



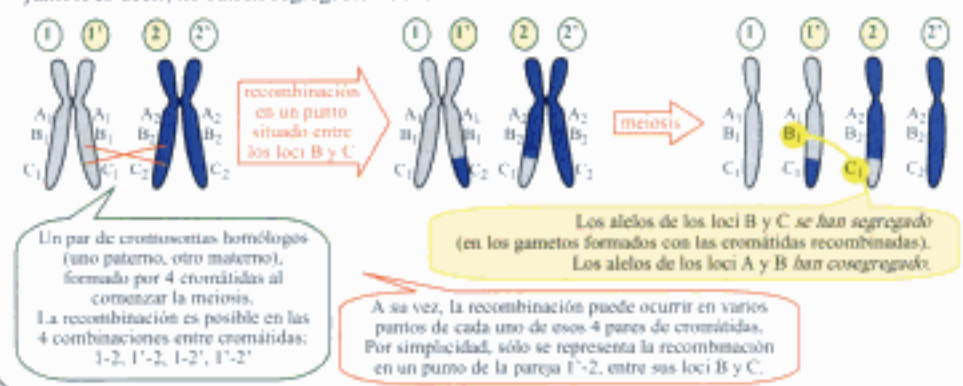
17.2.1 Fracción de recombinación como medida de distancia en un mapa genético

Durante la meiosis tiene lugar el **sobrecruzamiento**, o intercambio recíproco de fragmentos de DNA entre cromosomas homólogos (*recombinación homóloga o meiótica*, pág. 106). Debe destacarse que la recombinación nunca tiene lugar entre cromosomas diferentes, sino únicamente entre los dos miembros de un par de cromosomas homólogos. Los mecanismos moleculares de la recombinación no se describen en este libro, por corresponder a un tratado de Genética. Aquí interesan, esencialmente, las consecuencias del proceso. Por un lado, que los cromosomas *recombinantes* que aparecen en cada gameto ya no son idénticos a los originales (paterno y materno) y, por otro, que la variación genética resultante depende, obviamente, de la posición del sobrecruzamiento en cada par de cromosomas.

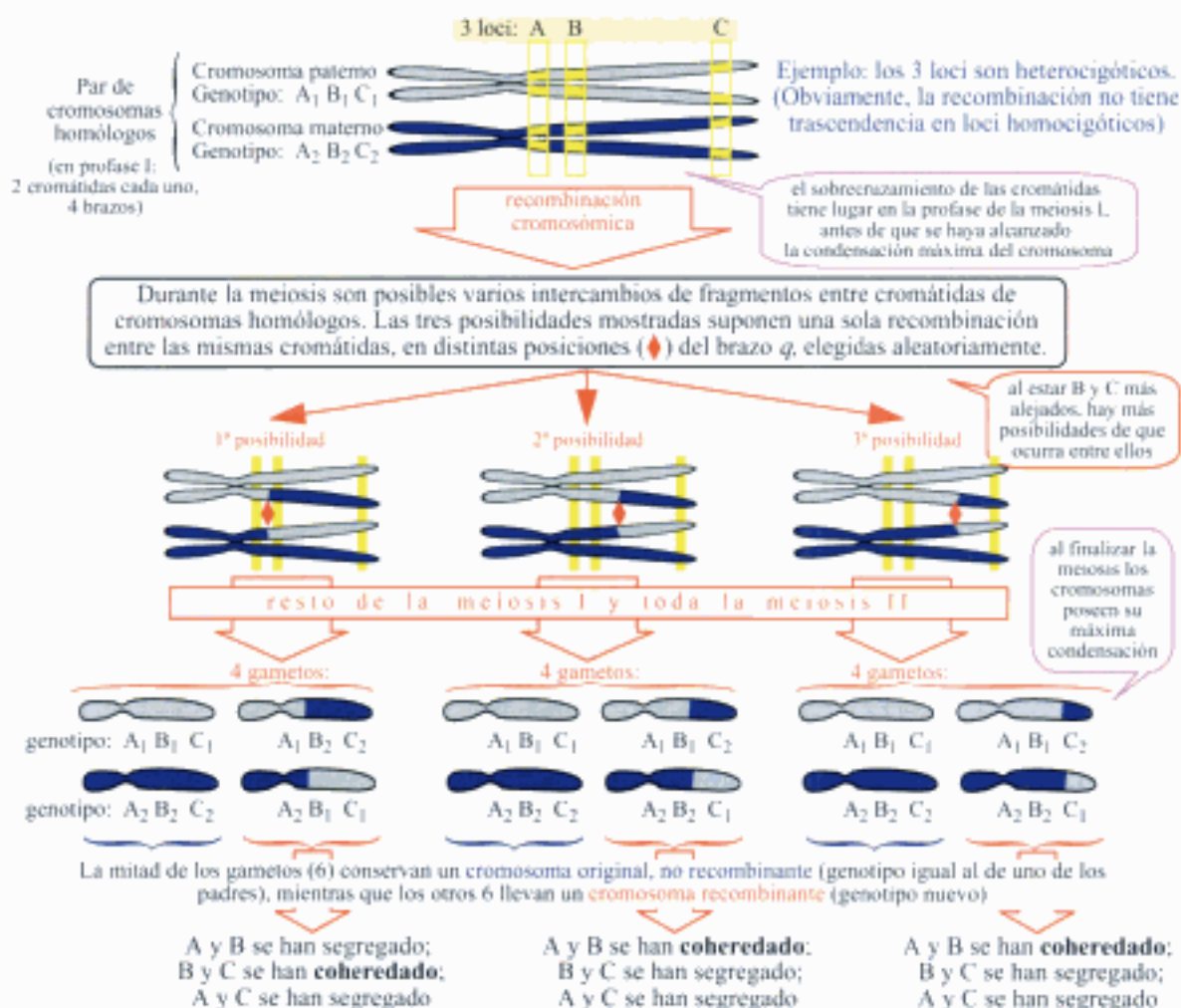
Como fase previa a la obtención de mapas genéticos, deben recordarse algunos términos básicos:

"Segregación" de cromosomas homólogos (y cromátidas) = **separación** de los cromosomas homólogos en la meiosis I (para formar parte de los gametocitos secundarios) y de las cromátidas en la meiosis II (para ir a parar a gametos diferentes). Cada par se segrega de forma **independiente** de los demás (pág. 105). Este proceso es independiente de que exista o no recombinación.

"Segregación" de alelos = **separación** de dos alelos de loci diferentes de un mismo cromosoma durante la meiosis (B y C). Sólo puede darse **si ha habido recombinación** entre esos loci, y se manifiesta en el momento de la segregación de los cromosomas y las cromátidas. El término "**co-segregación**" de alelos (A y B) puede parecer contradictorio, pues indica que éstos permanecen juntos, es decir, no sufren segregación entre sí.



La característica que define el mapeo genético es el análisis de la **fracción o frecuencia de recombinación**, definida como la probabilidad de que los alelos presentes en dos loci diferentes (correspondientes a la posición de los marcadores considerados) se separen o segreguen como consecuencia de la recombinación meiótica. Dicho de otro modo, es la frecuencia con la que dos alelos determinados aparecen recombinados en la descendencia. Como se discute a continuación, esta fracción depende de la **distancia entre los dos loci** en el cromosoma, lo que permite emplearla para establecer el mapa.



Conclusiones de interés, relacionadas con la construcción de mapas genéticos:

a) Si dos loci están muy próximos en el cromosoma (por ej., A y B):

Es difícil que se produzca recombinación en una posición intermedia entre ambos; por tanto, hay **baja frecuencia de recombinación** (2 de un total de 12 gametos = 16,7%).

Con alta probabilidad, los dos alelos quedan en un mismo cromosoma y cromátida: *se coheredan*.

Los alelos que se coheredan con frecuencia máxima (es decir, que casi nunca se segregan) se dice que están **ligados**. Lógicamente, los loci de cromosomas diferentes nunca están ligados, o tienen un **ligamiento génico** nulo.

Al conjunto de alelos localizados en una región pequeña del cromosoma se le denomina **haplotipo**. Se caracteriza porque sus alelos se transmiten conjuntamente (son alelos ligados).

b) A **mayor separación entre los loci** en el cromosoma (por ej., B y C o, más aún, A y C):

Existe mayor probabilidad de recombinación en posiciones intermedias; por tanto, hay **mayor frecuencia de recombinación** (para B y C: 4 de 12 = 33,4%; para A y C: 6 de 12 = 50%).

Al sufrir la recombinación, los dos alelos quedan en distinto cromosoma, por lo que hay

menor probabilidad de que se hereden juntos.

Nota: la terminología es aún más variada de la que se ha empleado aquí; por ejemplo, una alta frecuencia de recombinación es sinónimo de herencia conjunta infrecuente, coherencia infrecuente, segregación de alelos frecuente y cosegregación infrecuente. Esto puede dificultar, a veces, la comprensión del significado de "frecuencia de recombinación".

17.2.2 Obtención del mapa genético

Una vez conocidas las posiciones relativas de un número elevado de marcadores genéticos, se puede estudiar la frecuencia de recombinación con respecto a ellos de cualquier nuevo gen o secuencia y así situar su locus en el "mapa" definido por aquéllos. De esa forma, se dice que se ha "mapeado" su posición en el cromosoma. Una vez localizado, el nuevo gen o secuencia pasa a engrosar el conjunto de marcadores que forman el mapa.

La escala en los mapas genéticos se mide en unidades **centimorgan (cM)**. 1 cM corresponde, por definición, a la distancia que separa dos loci que experimentan recombinación en el 1% de las meiosis. Empíricamente, en el genoma humano esta distancia corresponde aproximadamente a 10^6 pb (1 Mb). A partir de la frecuencia de recombinación se deducen las distancias relativas entre varios loci y se puede establecer su ordenación en el cromosoma, comenzando así un mapa genético:

Como ejemplo muy simple, al estudiar tres loci cualesquiera, X, Y y Z:

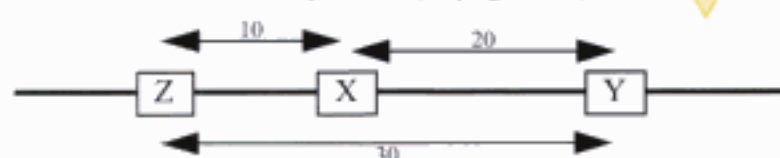
Frecuencia de recombinación observada para cada pareja de loci:

	X	Y	Z
X		20%	10%
Y	20%		30%
Z	10%	30%	

Distancia entre los loci:

	X	Y	Z
X		20 cM	10 cM
Y	20 cM		30 cM
Z	10 cM	30 cM	

interpretación (mapa genético):



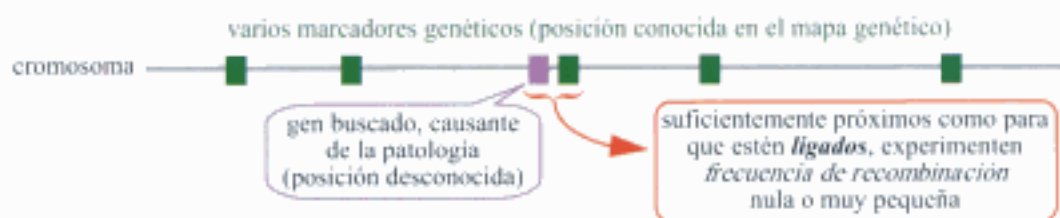
Se concluye que el locus X debe estar entre los loci Y y Z, o lo que es lo mismo, los loci están en el orden Z - X - Y (o viceversa, Y - X - Z).

Experimentalmente, los mapas genéticos requieren estudios familiares con árboles genealógicos, en los que se examina, para alguno de los marcadores, la combinación de alelos que va siendo heredada. Éstos se analizan en los distintos individuos a lo largo de varias generaciones; en humanos normalmente se parte de muestras de células sanguíneas nucleadas. La presencia de una u otra forma alélica se detecta por los procedimientos, variados, que permitan distinguirlas: bien a través de su influencia sobre el fenotipo (caracteres observables, aparición de enfermedad, etc.) o bien mediante técnicas moleculares de detección de secuencias (hibridación *in situ* por fluorescencia, método Southern o amplificación por PCR). A partir de la información de la herencia de alelos se calculan las frecuencias de

recombinación entre éstos. Con este planteamiento se han podido determinar, entre otros, los mapas genéticos de plantas, levadura, *Drosophila* y ratón.

En humanos es difícil determinar mediante estudios familiares las frecuencias de recombinación de loci génicos. Esto se debe, por un lado, a la escasez de casos en los que se transmiten simultáneamente dos alelos representativos, por ejemplo aquéllos causantes de dos enfermedades. Por otra parte, la baja tasa de reproducción del ser humano dificulta el cálculo fiable de dicha frecuencia. Para superar esta limitación, en la práctica se acude a **marcadores polimórficos**, es decir, aquéllos que presentan en la población un elevado número de alelos (formas alternativas), con lo cual son muy frecuentes los individuos heterocigóticos y la recombinación se podrá detectar en un alto porcentaje de la descendencia (recuérdese que la recombinación de loci homocigóticos no se puede apreciar). Los marcadores polimórficos son en general no génicos, no relacionados con enfermedades o con genes, sino simplemente secuencias de DNA detectables en una posición fija del cromosoma. Entre ellos destacan los basados en el DNA mini- y microsatélite que presentan polimorfismo en el número de repeticiones (VNTR, pág. 383). El número de marcadores polimórficos, especialmente los microsatélites, ha crecido de forma exponencial a lo largo de los años, de modo que actualmente se dispone de un mapa para cada cromosoma con la ubicación de un elevado número de marcadores.

La utilidad del mapa genético se ilustra, por ejemplo, con la posibilidad de localizar el gen responsable de una enfermedad hereditaria, aun sin conocer la identidad de dicho gen ni la base molecular de la enfermedad. Para ello, se busca, entre todos los que componen el mapa, un alelo de un marcador que esté presente en los individuos afectados pero no en los sanos, lo que indica que el locus de tal marcador está ligado al del gen buscado. Por consiguiente, el gen debe estar muy próximo a la posición conocida del marcador en el mapa. Esta posibilidad de ubicación de genes se potencia si se elabora el mapa con marcadores dispuestos a lo largo de todo el genoma; se asegura así, por puro azar, que siempre exista algún marcador suficientemente próximo al gen como para que ambos estén ligados. De esta forma se ha podido encontrar, por ejemplo, la localización cromosómica de los genes responsables de la fibrosis quística, la anemia falciforme, la enfermedad de Tay-Sachs, el síndrome X-frágil y la distrofia miotónica.



17.3 MAPAS FÍSICOS

Como se ha indicado, estos mapas miden directamente distancias físicas (es decir, en pares de bases) entre genes u otras regiones del DNA en el genoma. Las estrategias metodológicas son muy diversas y conducen a distintas variantes de mapas y niveles de resolución. Dependiendo de la escala o nivel de resolución, conocida como "unidad cartográfica o de mapa", se distinguen tres estrategias: baja resolución (varias Mb), alta resolución (hasta 10^3 bases) y máxima resolución (secuencia nucleotídica). En este capítulo se describen únicamente los dos primeros niveles, dejando la secuenciación del DNA para el tema siguiente.

17.3.1 Obtención de mapas físicos a baja resolución

Este nivel de mapeo corresponde a la asignación de cada marcador a un cromosoma o región cromosómica. Se utilizan distintos planteamientos que además implican muestras diferentes, según se describe a continuación. En general, el marcador se detecta empleando sondas o cebadores de PCR, ambos específicos para su secuencia; en el caso de un marcador génico, puede acudir también a la detección de su producto funcional (proteína o RNA).

17.3.1.1 Métodos basados en empleo de líneas celulares somáticas híbridas

Consisten en la búsqueda del marcador (génico o no) en líneas celulares híbridas, preparadas de tal forma que contienen material genético de dos orígenes diferentes: el genoma completo de un roedor (aunque con una mutación útil, explicada más adelante) y una parte del genoma humano, que puede ser varios cromosomas, uno solo o un fragmento de cromosoma. Disponiendo de una colección (o *panel*) de tales clones celulares híbridos, cada uno con partes distintas del genoma humano, se puede ensayar la presencia en todos ellos del marcador y así ubicar en qué cromosoma o parte del mismo se encuentra. Repitiendo el proceso para distintos marcadores se llega a construir un mapa físico.

Ejemplo de mapeo físico de un DNA de interés en cromosomas humanos. Para facilitar el análisis, se simplifica el ejemplo a 4 clones que sólo han retenido los cromosomas 1-5

		clon 1	clon 2	clon 3	clon 4
determinación del cariotipo humano presente en cada clon	cromosoma 1	no	si	no	no
	cromosoma 2	si	no	no	no
	cromosoma 3	no	no	si	si
	cromosoma 4	no	si	si	no
	cromosoma 5	no	no	no	si
	resto de cromosomas	no	no	no	no
resultado de la detección sobre cada clon (actividad enzimática, hibridación, etc.)		-	+	+	-

Conclusión:

la región de DNA que se está detectando (gen o DNA de interés) debe estar en el cromosoma 4

Selección de las células híbridas

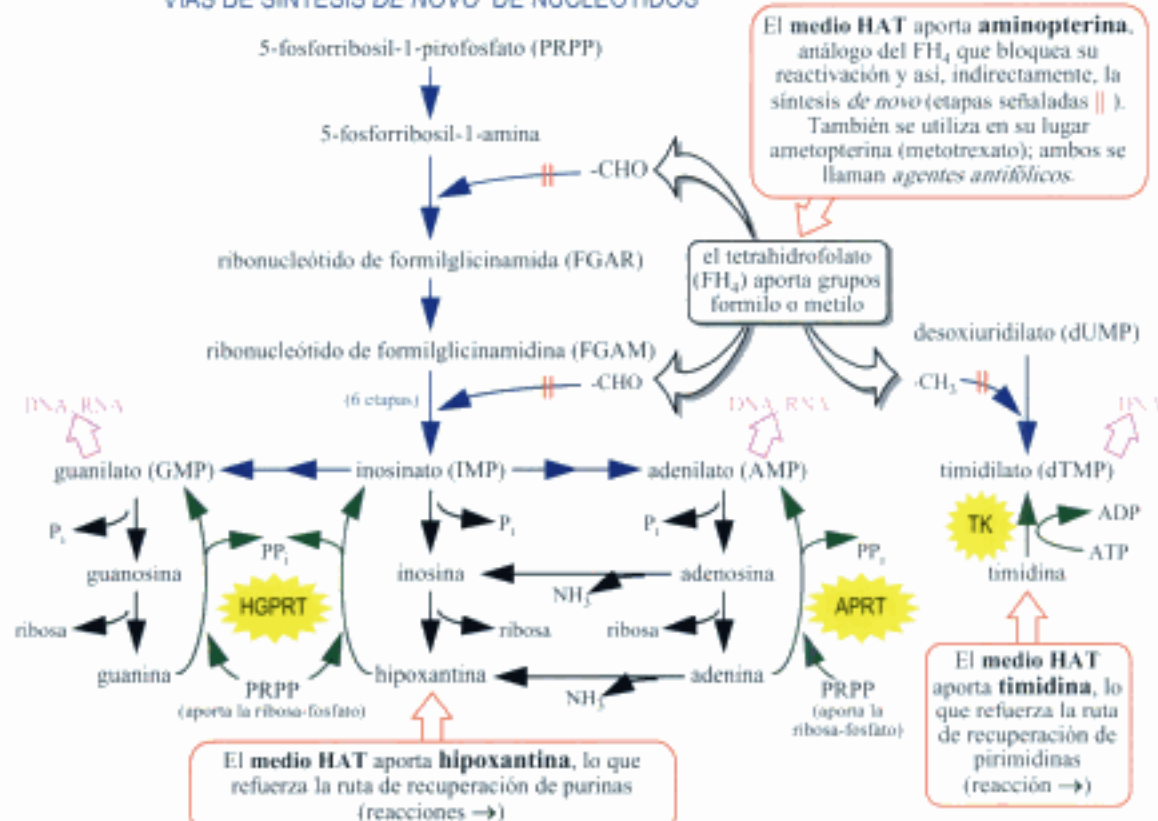
Un aspecto sumamente importante para el mapeo mediante células somáticas híbridas (que también se emplea en la preparación de anticuerpos monoclonales) es la capacidad de seleccionar los híbridos, es decir, que en cultivo sólo sobrevivan estas células. El sistema de selección más utilizado hace uso del **medio HAT**, así llamado por contener **Hipoxantina** (una base nitrogenada, 6-oxopurina), **Aminopterina** (un agente antifólico) y **Timidina** (un nucleósido).

1

Fundamento metabólico del uso de medio HAT

Su misión es **bloquear** la principal ruta de síntesis de nucleótidos, la denominada **via de novo**, al tiempo que **favorece** la ruta alternativa, **via de recuperación**.

VÍAS DE SÍNTESIS DE NOVO DE NUCLEÓTIDOS



VÍAS DE RECUPERACIÓN DE BASES PARA OBTENER NUCLEÓTIDOS

Y REACCIONES DE DEGRADACIÓN DE LOS NUCLEÓTIDOS, QUE APOYAN LAS VÍAS DE RECUPERACIÓN



Enzimas clave en las vías de recuperación:

HGPRT = hipoxantina/guanina fosforribosiltransferasa

APRT = adenina fosforribosiltransferasa

TK = timidina quinasa

2

Consecuencias del cultivo de las células en medio HAT

Para que el medio HAT sea eficaz se requiere la fusión de células humanas se realice con líneas de roedor mutantes, **deficientes en la actividad enzimática HGPRT**, esencial para la vía de recuperación. Las células de roedor dependen, pues, de la vía *de novo*, que no es usada habitualmente por las células normales. De esta forma, en el **híbrido** se complementa la capacidad de proliferación propia de la célula de roedor con la capacidad de crecimiento en medio HAT aportada por la presencia de material genético humano (que sí posee el gen de la enzima HGPRT normal).

Tipo de célula	Ruta metabólica	Utilización de las rutas metabólicas y, como consecuencia, crecimiento en cultivo	
		en medio normal	en medio HAT
Célula humana o de roedor normales  No prolifera bien en cultivo. Posee actividad HGPRT.	Vía de novo → Vía de recuperación →	sí sí no (característica del tipo celular) 	no, bloqueada por A sí, favorecida por H y T no (característica del tipo celular) 
Línea celular de roedor especial para fusión  Prolifera fácilmente en cultivo. Mutante HGPRT (carece de esa actividad enzimática).	Vía de novo → Vía de recuperación →	sí no, por falta de HGPRT sí (característica del tipo celular) 	no, bloqueada por A no, por falta de HGPRT no , por falta de nucleótidos 
Célula híbrida resultante de la fusión  Combina la capacidad de proliferar en cultivo de la de roedor y la actividad enzimática HGPRT de la humana	Vía de novo → Vía de recuperación →	sí sí sí 	no, bloqueada por A sí, favorecida por H y T sí 
Conclusiones:		El medio normal no permite la selección	Gracias al medio HAT se seleccionan las células híbridas

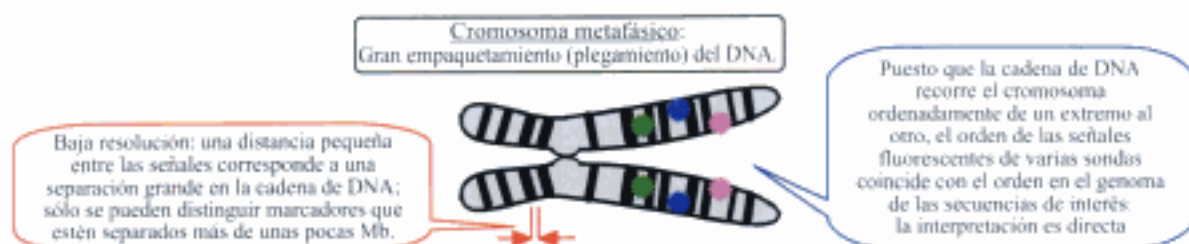
La preparación de híbridos a partir de células somáticas humanas sólo permite identificar el cromosoma en el que está situado el DNA de interés. Se puede alcanzar mayor resolución de mapeo empleando híbridos similares preparados con una porción menor del genoma humano: *híbridos monocromosómicos*, conteniendo un solo cromosoma, e *híbridos subcromosómicos*, con un fragmento de cromosoma. Con estos últimos se construye lo que se denomina "mapa de asignación regional", es decir, la localización de cada marcador en una región determinada de un cromosoma.

17.3.1.2 Obtención de mapas mediante ensayos de hibridación con sondas fluorescentes

La ubicación de marcadores para formar un mapa físico de baja resolución puede también conseguirse realizando ensayos de hibridación con sondas que contengan la propia secuencia del marcador. Esta técnica ha adquirido gran aplicabilidad en el mapeo gracias a las sondas fluorescentes (**hibridación in situ por fluorescencia** o **FISH**, pág. 172). La sonda se marca con un fluorocromo, bien de forma directa o a través de un grupo "indicador" (*reporter*, por ejemplo, biotina o digoxigenina, pág. 185). Para identificar el cromosoma o la región del mismo donde hibrida la sonda se puede acudir a la observación bajo el microscopio de fluorescencia o a la citometría de flujo.

a) Mapas por hibridación in situ (FISH) cromosómica

La hibridación se realiza sobre preparaciones para el microscopio de cromosomas metafásicos, previamente desnaturalizados. La identidad de los cromosomas se revela por las técnicas habituales de establecimiento del cariotipo (forma, tamaño y bandas bajo tinción). El DNA diana (marcador) se manifiesta bajo el microscopio de fluorescencia como puntos luminosos dobles, debido a la hibridación de la sonda con las dos cromátidas. Se alcanza con esta técnica un nivel de resolución cromosómico o subcromosómico, y es posible localizar o mapear simultáneamente varios marcadores empleando varias sondas con fluorocromos distintos.



b) Mapas por citometría de flujo

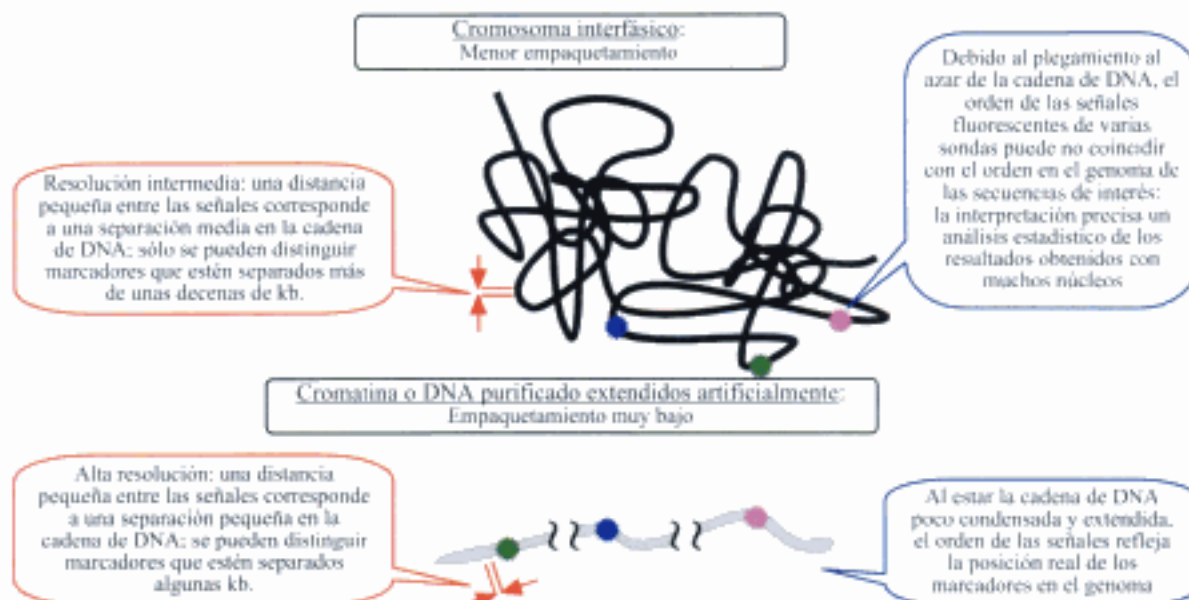
La identificación del cromosoma en el que se encuentra cada marcador puede también realizarse si la sonda se hibrida con muestras que contienen cada una un solo tipo de cromosoma. Éstas pueden conseguirse aplicando la clasificación de células activada por fluorescencia (FACS, pág. 126), adaptada al mapeo del modo siguiente: Los cromosomas se tratan con fluorocromos que se intercalan entre las bases nitrogenadas apiladas del DNA; los de mayor longitud pueden incorporar mayor cantidad de ese agente intercalante, por lo que adquieren más fluorescencia. Esto permite diferenciar los cromosomas en la etapa de análisis o caracterización, proporcionando lo que se ha llamado un *cariograma de flujo*, así como separarlos en la etapa de clasificación y obtener muestras, cada una con un solo tipo de cromosoma. Finalmente, estos cromosomas separados pueden emplearse, por ejemplo, en hibridación *dot-blot* con la sonda específica para el marcador buscado.

17.3.2 Mapas físicos de alta resolución

Estos métodos, complementarios a los anteriores, permiten alcanzar una resolución inferior a 1 Mb, incluso hasta llegar a unos pocos nucleótidos, lo que ya supone el **nivel molecular** de resolución; de ahí que sean los métodos de mapeo físico más importantes. Sirven de enlace entre los mapas físicos anteriores y los de secuenciación o mapas de resolución total a escala de un nucleótido (tema 18). Se plantean principalmente dos tipos de abordaje: hibridación y enzimas de restricción.

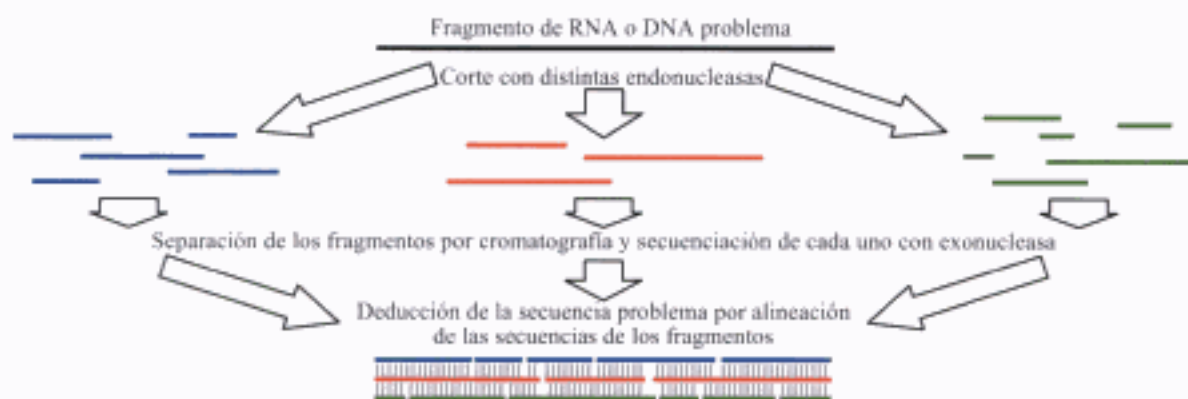
17.3.2.1 Mapa físico por hibridación *in situ* (FISH) de alta resolución

La baja resolución asequible sobre cromosomas metafásicos se puede aumentar mediante la hibridación sobre cromosomas menos compactados.



17.3.2.2 Mapa físico con enzimas de restricción

Antes de que se descubrieran las endonucleasas de restricción y se extendiera su uso en el laboratorio de investigación, se utilizaron **otras nucleasas** para secuenciar RNAs, siguiendo el mismo enfoque utilizado previamente para secuenciar polipéptidos, es decir, la fragmentación de la molécula por puntos de características conocidas y la reconstrucción del conjunto mediante solapamiento de los fragmentos. Este método no dio los resultados esperados con DNAs, debido a su mayor longitud y, sobre todo, a la menor especificidad de las DNAsas, por lo que sólo se consiguió secuenciar algunos oligodesoxirribonucleótidos.



La ventaja de estos planteamientos iniciales es que sirvieron de pauta para la confección de los actuales mapas físicos del genoma a partir de fragmentos de DNA originados por **enzimas de restricción**. Un mapa de restricción para una región determinada del genoma está constituido por una relación ordenada de sitios que pueden ser cortados por enzimas de restricción concretas, con indicación de las distancias entre ellos (en pb). A diferencia de los mapas genéticos, para su elaboración no afecta la funcionalidad de la región de DNA (su carácter génico) ni la presencia en ella de mutaciones (salvo que éstas alteren precisamente el sitio de restricción). Por el contrario, si importan algunas alteraciones que puede sufrir el DNA, como translocaciones, deleciones o inserciones.

La resolución o escala de los mapas de restricción depende de la separación entre los sitios de corte o, dicho de otro modo, de la frecuencia con que la secuencia diana reconocida por la enzima aparece en el genoma (*frecuencia de reconocimiento del genoma por la enzima*). Por ejemplo, asumiendo que en un DNA de gran longitud las cuatro bases queden distribuidas aleatoriamente, cuanto más corta sea la secuencia del sitio de restricción, más probabilidad habrá de encontrarla, por puro azar, es decir, mayor será la frecuencia de reconocimiento:

tamaño de la secuencia diana de restricción	estadísticamente, aparecería una diana cada:	tamaño teórico medio de los fragmentos	nº teórico de dianas en el genoma humano haploide
4 pb	$4^4 = 256$ pb	256 pb	$13 \cdot 10^6$
6 pb	$4^6 = 4.096$ pb	4.096 pb	$8 \cdot 10^5$
8 pb	$4^8 = 65.536$ pb	65.536 pb	$5 \cdot 10^4$

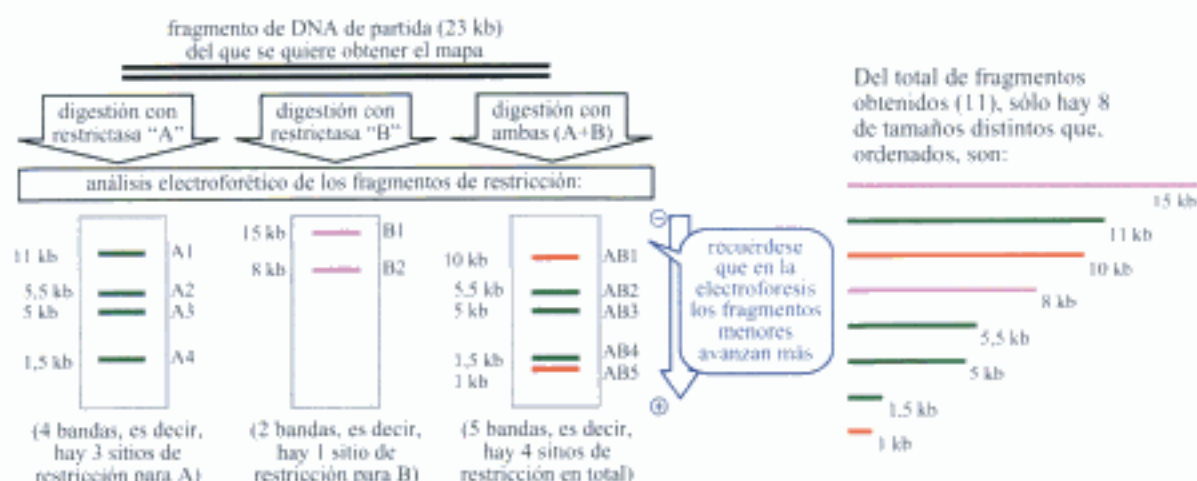
La frecuencia suele ser inferior a lo teóricamente calculado, debido a varias causas. Por un lado, la distribución de bases en el genoma no es totalmente aleatoria, lo que conduce a fragmentos de longitud muy variable (frente a las cifras promedio indicadas). Como ejemplo, los fragmentos de DNA humano generados por distintas enzimas de restricción varían entre los 300 pb en promedio obtenidos con *Alu I* (diana de 4 pb) y las 9,7 Mb para *Not I* (diana de 8 pb). Otra causa que reduce el número de sitios de corte en el genoma es la presencia de regiones ricas en secuencias CpG, que al estar normalmente metiladas no son sensibles a las enzimas de restricción.

La construcción de un mapa de restricción se simplifica si el número de sitios de restricción es reducido. Esto es especialmente necesario para mapear moléculas grandes de DNA, en las que las secuencias específicas para las enzimas de restricción habituales se hacen muy numerosas. Para estos fragmentos grandes el mapeo sólo se puede abordar empleando enzimas "cortadoras infrecuentes" y electroforesis en campo pulsante. En cualquiera de los casos, el revelado de los geles y la hibridación Southern con diversas sondas permiten deducir la distribución de los sitios de restricción, o "cartografía" del mapa.

Para facilitar la comprensión del método de obtención de mapas físicos de restricción, se proponen a continuación unos ejemplos prácticos. En el primer caso, se emplean dos enzimas (A y B), actuando aisladamente o en forma conjunta (experimento de doble digestión). En el segundo ejemplo, las dos enzimas actúan de forma sucesiva. En el tercero, la digestión enzimática es idéntica al primero, pero se emplea la hibridación para enriquecer la información obtenida.



Ejemplo de aplicación de las enzimas de restricción para investigar la posición de determinadas secuencias de DNA (mapeo físico), dentro de un fragmento de genoma de interés (por ejemplo, por la presencia de un gen)

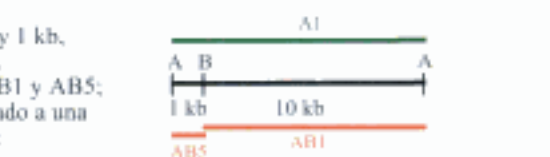


Resultados:

- En las 3 digestiones, la suma de los tamaños de las bandas es de 23 kb (igual al fragmento inicial)
- Sólo 3 bandas de la digestión con A coinciden en la digestión doble
- Las 2 bandas de la digestión con B no aparecen en la digestión doble

Interpretación:

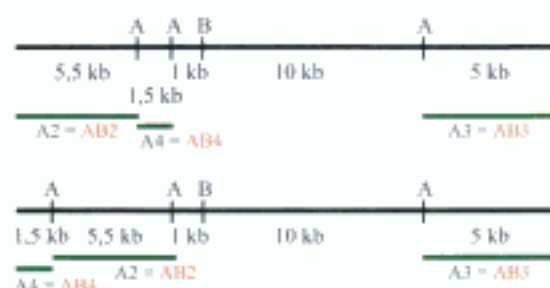
- En la digestión doble aparecen 2 bandas nuevas, de 10 y 1 kb, cuyos tamaños suman el de la banda A1 desaparecida.
Conclusión: el fragmento A1 es escindido por B en AB1 y AB5; es decir, hay un sitio de restricción B intermedio, situado a una distancia de 1 kb de un sitio A y 10 kb de otro sitio A:



- Los 2 fragmentos de B (8 y 15 kb) pueden estar situados sólo de una forma, a ambos lados del punto de corte B anterior:



- La asignación de los fragmentos A2, A3 y A4 (iguales a AB2, AB3 y AB4) es la siguiente: La región de 7 kb en el mapa anterior sólo puede corresponder a la suma de A2 y A4. La región de 5 kb debe corresponder a A3. Dependiendo del orden de colocación de estos 2 fragmentos, surgen dos posibles soluciones, que no pueden discernirse con los resultados obtenidos:



Conclusión final:

Mapa de restricción

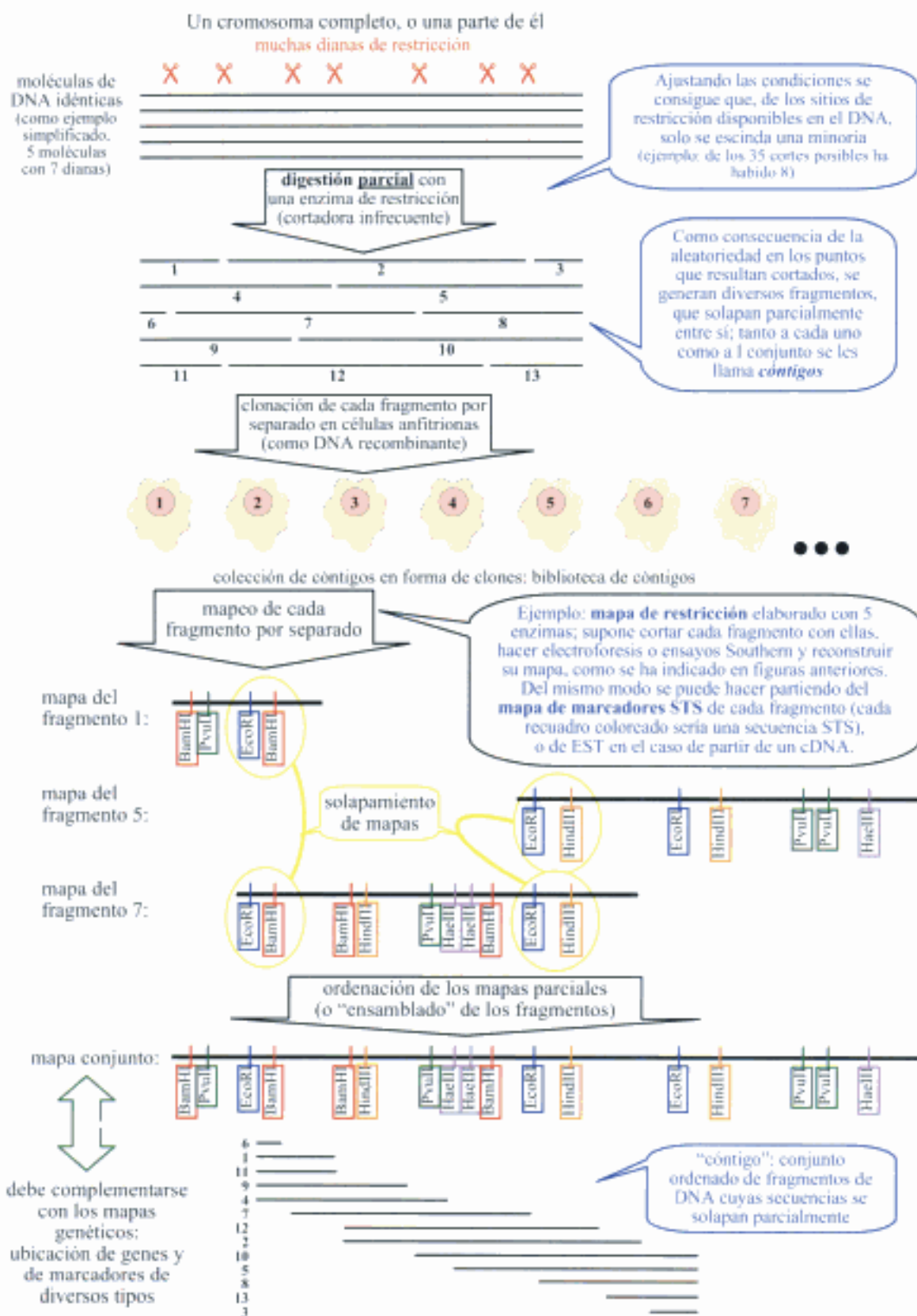


(obviamente, la misma molécula puede representarse de derecha a izquierda, con lo que los valores que figuran en la escala serían otros pero las distancias son las mismas: 0 5 15 16 17,5 23 y 0 5 15 16 21,5 23 respectivamente)

A B A A A B A A

17.3.2.4 Mapas de cóntigos

Se denomina así a los mapas que se obtienen mediante el solapamiento de fragmentos de DNA. Estos fragmentos proceden generalmente de la digestión parcial con enzimas de restricción de otros fragmentos mayores, cada uno de los cuales se ha obtenido a su vez por fragmentación de los 24 tipos diferentes de cromosomas (en humanos) de la muestra. El planteamiento del mapeo es idéntico en ambos niveles, y similar al descrito para las nucleasas. Para la deducción del orden o solapamiento de los diferentes fragmentos se acude a la comparación de los marcadores presentes en cada uno, que constituyen generalmente mapas de restricción o mapas de STSs o ESTs.



Tema 18

Secuenciación del genoma y Proyectos Genoma

18.1	Introducción	231
18.2	Método químico de secuenciación del DNA	231
18.2.1	Marcaje de una hebra del DNA	232
18.2.2	Modificación de las bases y fragmentación del DNA	232
18.2.3	Separación de los fragmentos y análisis	234
18.3	Método enzimático de secuenciación del DNA	235
18.3.1	Síntesis de un DNA complementario de longitud variable	235
18.3.2	Separación de los fragmentos y análisis	237
18.3.3	Automatización	237
18.3.4	Variantes del método	238
18.4	Secuenciación del RNA	238
18.5	Proyecto Genoma Humano (PGH): Características esenciales	239
18.6	Otros proyectos de secuenciación del genoma humano	240
18.7	Genómica comparativa	241

18.1 INTRODUCCIÓN

Los esfuerzos iniciales para secuenciar ácidos nucleicos dieron pocos resultados. Hasta los años 70 era difícil y costoso secuenciar incluso fragmentos de sólo 5 a 10 nucleótidos. Fueron los métodos químico de Maxam y Gilbert (1977) y enzimático de Sanger (1980) los que abrieron el camino a dicho gran objetivo. Se consiguieron moléculas mayores de DNA y se definieron algunas secuencias concretas como las del fago ϕ X174 (5,4 kb), mitocondria humana (16,5 kb), fago lambda (48,5 kb) y virus animal de Epstein-Barr (170 kb).

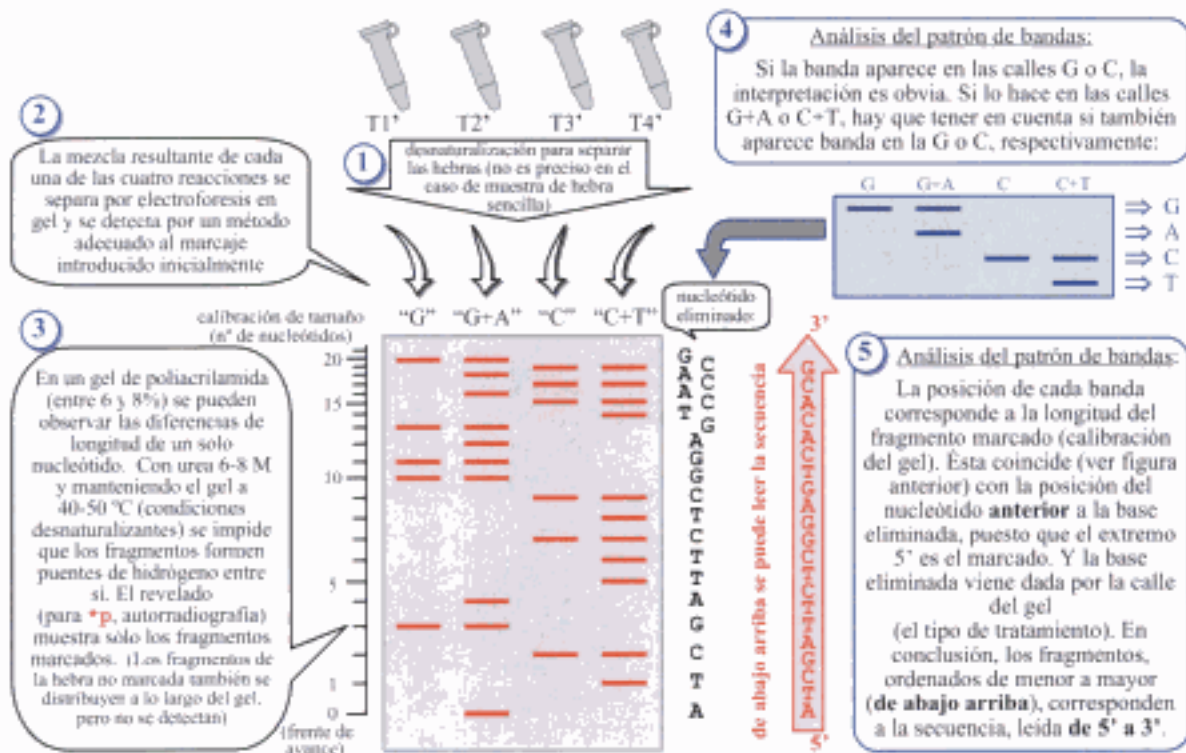
Estos dos métodos pioneros de la secuenciación tienen en común algunos aspectos, como la utilización de fragmentos del DNA original escindidos con enzimas de restricción, el marcaje del DNA de forma radiactiva o química (compuestos fluorescentes), el empleo de métodos electroforéticos para separar los fragmentos generados, etc. La muestra de DNA de partida para la secuenciación procede generalmente de clonación celular, de amplificación por PCR o de la digestión controlada con una enzima de restricción de un genoma, un cromosoma u otro fragmento de DNA. Con frecuencia, las muestras se amplifican por PCR para disponer de suficiente cantidad. Igualmente, es conveniente la purificación para eliminar DNA contaminante en la muestra (por ej. DNA bacteriano) y reactivos procedentes de la purificación o de otros procesos previos (como cebadores y nucleótidos restantes de la PCR). Se puede acudir también a recuperar del gel el DNA contenido en una banda de electroforesis.

18.2 MÉTODO QUÍMICO DE SECUENCIACIÓN DEL DNA

Este procedimiento, uno de los más usados pero actualmente desplazado por el método enzimático, se basa en la hidrólisis química y en un diseño original aplicable a secuencias de DNA de menos de 250 nucleótidos. Se conoce habitualmente como **método de Maxam y Gilbert**. Puede describirse en tres etapas: marcaje del DNA, hidrólisis química selectiva del DNA y análisis de los productos.

18.2.3 Separación de los fragmentos y análisis

Para analizar las 4 mezclas de fragmentos obtenidas, se separan todos ellos de acuerdo con su tamaño, empleando electroforesis. El tamaño de cada fragmento marcado indica directamente la posición en la secuencia del nucleótido correspondiente al corte.



18.3 MÉTODO ENZIMÁTICO DE SECUENCIACIÓN DEL DNA

El método enzimático o **de Sanger** es la base de las variantes más utilizadas actualmente para la secuenciación a gran escala del DNA. En él no se degrada el DNA como en el método químico, sino que se acude a la interrupción controlada de la síntesis de una hebra complementaria durante una replicación *in vitro*. Esta síntesis, catalizada por una DNA polimerasa, define el carácter enzimático del método.

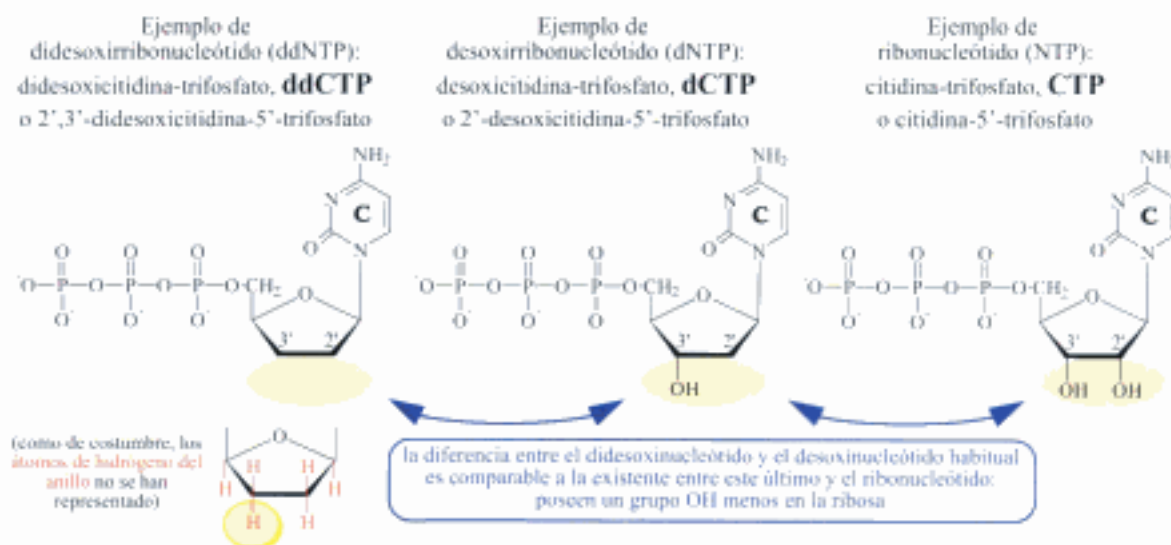
Puesto que la actividad 5'-exonucleasa de las DNA polimerasas naturales no es conveniente aquí, se suelen usar variantes, tales como el denominado "fragmento de Klenow" o fragmento grande de la DNA-polimerasa I de *E. coli* (pág. 147), o formas alteradas artificialmente de la DNA polimerasa del fago T7 (algunas comercializadas con el nombre de "secuencasa", Sequenase®). También se emplean en ocasiones transcriptasas inversas o la DNA polimerasa Taq (polimerasa termoestable de la arqueobacteria *Thermus aquaticus*, conocida por su aplicación en PCR, pág. 189). Estas alternativas intentan, además de evitar la actividad 5'-exonucleasa, conseguir mayores procesividad, actividad enzimática y eficacia sobre regiones homopoliméricas y regiones con tendencia a adoptar estructura secundaria. Todo ello conduce a la posibilidad de secuenciar cada vez fragmentos más largos y con menor margen de error.

Igualmente, han ido surgiendo distintas variantes del método inicial en lo relativo al tipo y lugar de marcaje de los nucleótidos, especialmente al aparecer los fluorocromos como alternativa a los isótopos radiactivos. Para simplificar, y puesto que el fundamento no varía, se expone a continuación una de ellas, que emplea cuatro fluorocromos distintos y ofrece la posibilidad de automatización. El método puede describirse en tres secciones: síntesis enzimática, análisis de los fragmentos y automatización. Al final se indicarán los matices que supone el empleo de variantes alternativas del método.

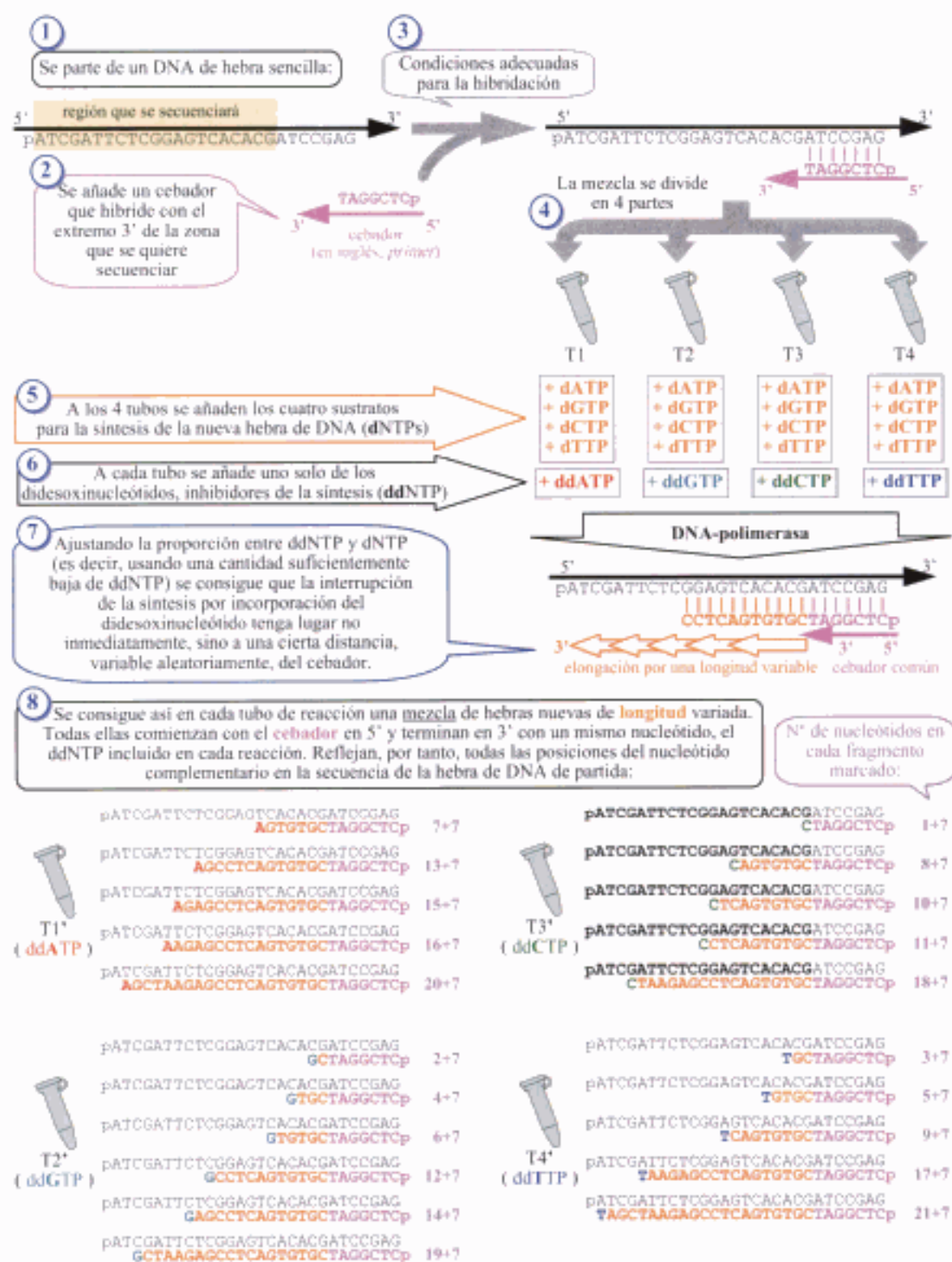
18.3.1 Síntesis de un DNA complementario de longitud variable

Dado que se emplea una DNA polimerasa para sintetizar hebras complementarias a la que se quiere secuenciar, una primera característica del método es la necesidad de un cebador, un oligonucleótido diseñado para que se hibride con el extremo 3' de la región que se quiere secuenciar.

La segunda, y principal, característica de este método de secuenciación es el uso de **didesoxinucleótidos**, análogos estructurales de los desoxinucleótidos que provocan la detención de la reacción de síntesis de DNA. Por ello, se conoce también como **secuenciación didesoxi** o **de terminación de cadena**.

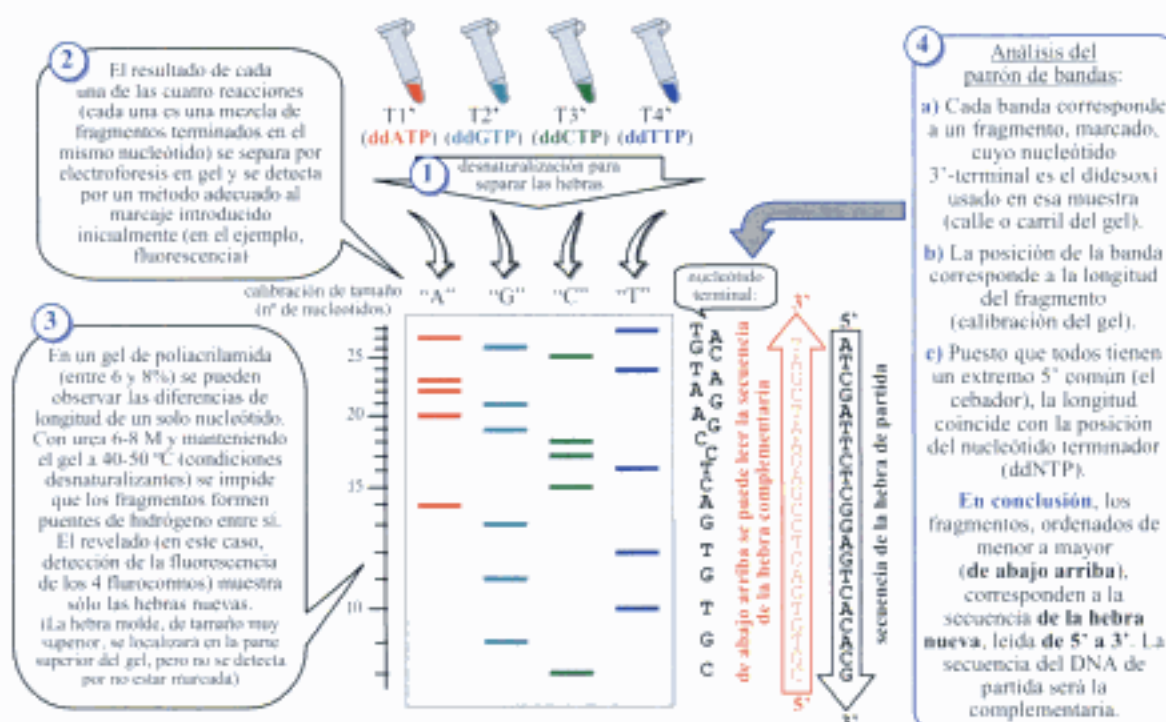


La razón para utilizar didesoxinucleótidos es que **compiten con los desoxinucleótidos** en la reacción de síntesis. La DNA polimerasa utiliza el **ddNTP** como sustrato, por su analogía estructural con el **dNTP**, y queda incorporado a la hebra en crecimiento como **ddNMP**, pero a partir de él la cadena no se puede elongar debido a la falta de grupo -OH en 3', que impide que forme enlace fosfodiéster con el siguiente nucleótido. Los ddNTPs, por tanto, **detienen la síntesis**.



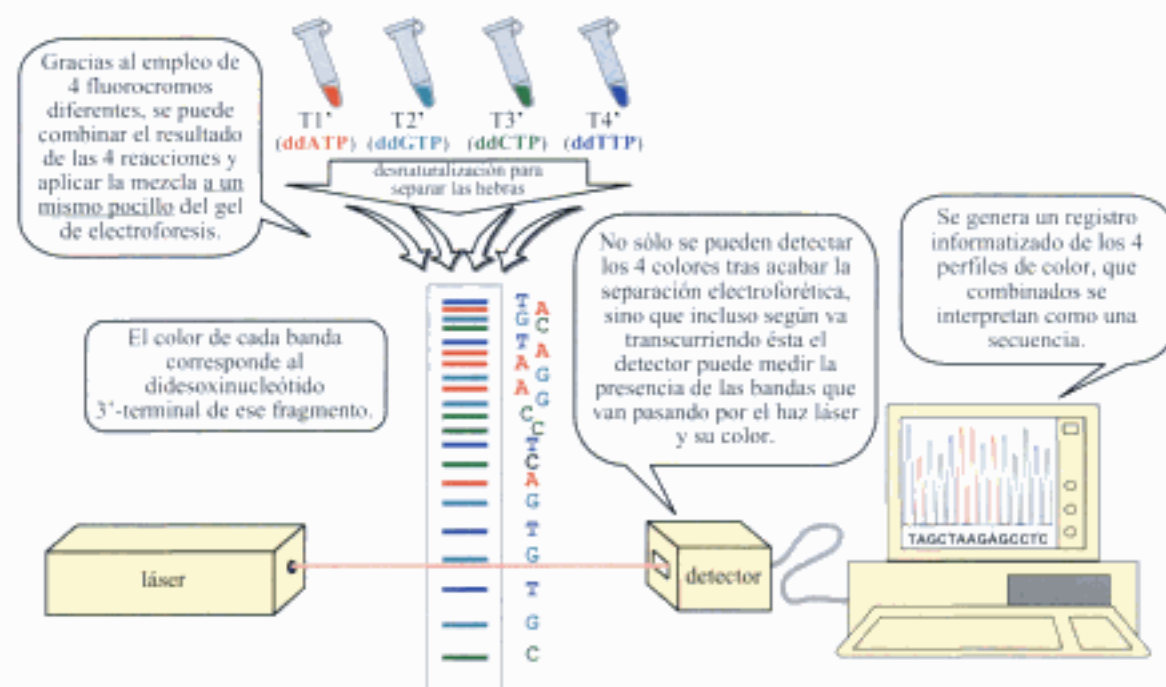
18.3.2 Separación de los fragmentos y análisis

Se sigue un procedimiento completamente análogo al análisis del método químico:



18.3.3 Automatización

La posibilidad de usar un fluorocromo distinto en cada una de las 4 reacciones de síntesis permite automatizar el método, de forma que se "lean" simultáneamente las hebras marcadas componentes de las 4 mezclas. Además, esta lectura de la fluorescencia se puede hacer en continuo, con lo que la electroforesis sigue transcurriendo, las moléculas de menor tamaño, ya detectadas, se salen por el extremo inferior del gel, y se sigue consiguiendo la separación de moléculas mayores en el gel, de modo que se amplía el número de nucleótidos que se pueden secuenciar en un solo experimento.



A título ilustrativo, de los numerosos fluorocromos disponibles, se indican los 4 usados más comúnmente para secuenciación en la actualidad (todos como N-hidroxysuccinimidilésteres, forma activa que permite su unión a un grupo amino del nucleótido).

Nombre común	nombre completo	propiedades de fluorescencia		
		λ_{ex} (nm)	λ_{em} (nm)	color emitido
FAM	(5 ó 6)-carboxifluoresceína	494	518	rojo anaranjado
JOE	(5 ó 6)-carboxi-4',5'-dicloro-2',7'-dimetoxifluoresceína	522	550	rojo-púrpura
TAMRA	(5 ó 6)-carboxitetrametilrodamina	555	580	púrpura
ROX	(5 ó 6)-carboxi-X-rodamina	580	605	azul verdoso

18.3.4 Variantes del método

Además de mediante el método ilustrado, el marcaje de la hebra de nueva síntesis se puede conseguir de otras maneras, aplicándose a otros componentes de la reacción en la que se sintetizan dichas hebras:

(los nº corresponden al esquema de la síntesis)	Método de marcaje de la hebra nueva							
	En el didesoxinucleótido (extremo 3')			En el cebador (extremo 5')			En los desoxinucleótidos (a lo largo de la hebra)	
	fluorescencia	radioactividad		fluorescencia	radioactividad		fluorescencia	radioactividad
2 cebador, con la misma secuencia en los 4 tubos (fracciones en que se divide la muestra)	sin marcar			marcado con fluorocromos distintos (nota 1)	marcado con el mismo fluorocromo (nota 1)	marcado con un isótopo (nota 1)	sin marcar	
3 4 dNTPs, en cada tubo los 4	sin marcar			sin marcar			3 sin marcar, 1 marcado (con 1 fluorocromo o 1 isótopo) (nota 2)	
4 4 ddNTPs, en cada tubo uno sólo	marcados con fluorocromos distintos	marcados con un mismo fluorocromo	marcados con un mismo isótopo	sin marcar			sin marcar	
detección de fluorescencia	✓	✓		✓	✓		✓	
autoradiografía			✓			✓		✓
electroforesis en 4 calles independientes y lectura tras terminar		✓	✓		✓	✓	✓	✓
electroforesis en 1 calle y lectura simultánea (posibilidad de automatización)	✓			✓				

corresponde al ejemplo ilustrado en las figuras

totalmente comparable al ejemplo ilustrado, salvo que la división de la muestra en 4 debe hacerse antes de añadir el cebador

método original de secuenciación

Nota 1: Empleando nucleótidos marcados durante su síntesis química o bien por modificación enzimática del cebador completo, mediante los métodos de marcaje terminal (véase el marcaje de sondas).

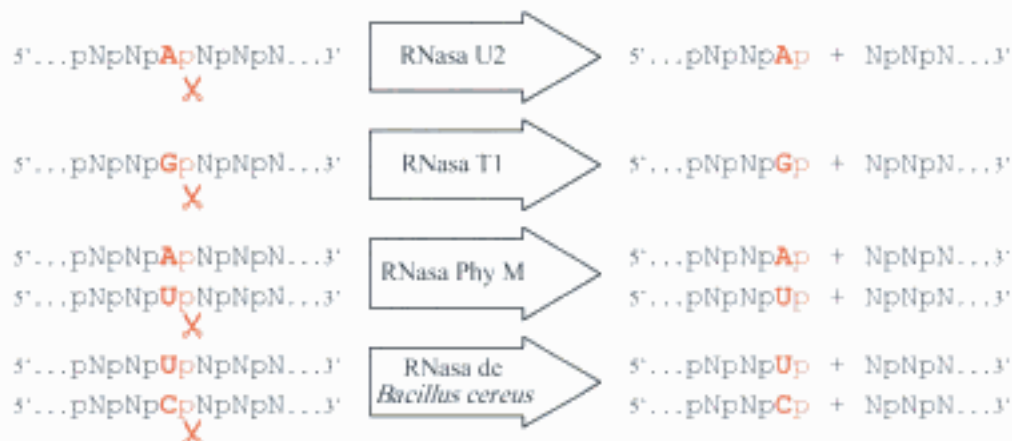
Nota 2: Basta con uno de los dNTPs marcado con isótopo radiactivo ($\alpha^{32}\text{P}$, $\alpha^{33}\text{P}$, $\alpha^{35}\text{S}$) o con fluorocromo, que se incorporará en varios puntos a la hebra en crecimiento. Nótese que, a diferencia del método químico de secuenciación, en este caso el isótopo debe estar en el fosfato α (no en el γ), que es el único que se conserva incorporado en el DNA.

Se han desarrollado además de estas modificaciones, otras que afectan al planteamiento y alcance del método químico y, especialmente, del enzimático, y que están siendo aplicadas en los distintos proyectos genoma bajo estudio en el momento actual. Sin embargo, creemos que su descripción excede las posibilidades de este texto. Baste citar entre ellas la secuenciación cíclica, la secuenciación múltiple, la quimioluminiscente y la secuenciación en fase sólida.

18.4 SECUENCIACIÓN DEL RNA

La secuenciación del RNA es, en general, más difícil que la del DNA, entre otras causas por su menor estabilidad. Es posible deducirla a partir de la secuencia del DNA del que se transcribe, o también sintetizar un cDNA a partir del RNA de interés y secuenciar aquél. Sin embargo, a veces es necesario secuenciar directamente el RNA, sobre todo cuando se desea determinar las posiciones de nucleósidos modificados, característicos de los tRNAs y rRNAs. Para

ello, se ha desarrollado un procedimiento que emplea distintas RNasas que cortan en el lado 3' de nucleótidos específicos, con lo que se puede obtener una serie de fragmentos que reflejan las posiciones de cada base. Siguiendo un planteamiento análogo al de la secuenciación química del DNA se puede entonces deducir la secuencia del RNA.



Asimismo, al menos en teoría, es posible aplicar el método de secuenciación dideoxi a RNAs. Lógicamente, en lugar de la DNA polimerasa usada para secuenciar el DNA, se precisa una transcriptasa inversa, capaz de emplear el RNA de interés como molde para sintetizar una hebra de DNA complementario (pág. 196), partiendo de la misma mezcla de dNTPs y de un ddNTP distinto en cada tubo.

18.5 PROYECTO GENOMA HUMANO (PGH): CARACTERÍSTICAS ESENCIALES

La consecución del mapa completo del genoma humano es un sueño del siglo XIX, que tomó forma en la segunda mitad del siglo XX con la introducción de los mapas genéticos y físicos, y que se completó con el inicio a partir de 1990 de la secuenciación, gracias a la puesta a punto y la automatización del método enzimático de Sanger.

Se trata de un enorme esfuerzo de investigación, bajo la dirección inicial de J. Watson, financiado por los Institutos Nacionales de Salud de EE.UU. y la Fundación Wellcome Trust del Reino Unido, con un ambicioso objetivo, la secuenciación en 15 años de los 3.300 millones de pares de bases del genoma humano, incluyendo, por tanto, la parte codificante (3% del total, formando unos 100.000 genes, aunque recientemente se dan cifras aún más discordantes, entre 30.000 y 200.000 genes) y la no codificante (el 97% restante). Han participado 16 centros públicos y privados de investigación de 6 países (Alemania, Canadá, EE.UU., Francia, Japón y Reino Unido). Entre aquéllos cabe citar, por ser responsables del 85% de la secuenciación, el Centro Sanger (Cambridge, Reino Unido), la Universidad de Washington (San Luis, Missouri, EE.UU.), el Joint Genome Institute (California, EE.UU.), el Instituto Whitehead (Cambridge, Massachusetts, EE.UU.) y la facultad de medicina Baylor (Houston, Texas, EE.UU.). El director actual del proyecto es Francis Collins y, en España (que no ha intervenido en este proyecto), Santiago Grisolia preside el comité para el PGH de la UNESCO. El Proyecto ha sido comparado, por su importancia, al descubrimiento de la Tabla Periódica de los Elementos Químicos (Mendeleiev) y, por su costo y complejidad, a la fabricación de la bomba atómica (Proyecto Manhattan) o la llegada del hombre a la Luna (Proyecto Apollo).

Experimentalmente, el PGH consta de las siguientes **grandes etapas**:

1. Se parte de muestras de DNA humano, de voluntarios, extraídas de sangre y mantenidas para su conservación en nitrógeno líquido (-196 °C).
2. El DNA se escinde en fragmentos, por puntos concretos, mediante enzimas de restricción. Esto se hace en dos fases, una para originar fragmentos relativamente grandes de unos 150 kb y, a partir de éstos, una segunda escisión para dar fragmentos más pequeños.
3. Se clonan los fragmentos, mediante la inserción de cada uno en un vector (plásmido, BAC, YAC ...) e incorporación a una célula anfitriona (bacteria o levadura). En junio de 2000 se habían clonado unos 24.000 fragmentos, equivalentes a un 97% del total del genoma.
4. Separación del DNA a partir de cada clon y, en su caso, amplificación por PCR, para obtener miles de copias idénticas de cada fragmento.
5. Secuenciación completa y repetida de cada fragmento de DNA (unas 7 veces en promedio), por el método de Sanger. En junio de 2000 se había secuenciado un 87% del genoma, con una precisión del 99%. Está prevista su finalización para el año 2003.
6. Ordenación de los fragmentos secuenciados para establecer su posición original en el genoma.

Para todas estas etapas, especialmente para la secuenciación y el ordenamiento de los fragmentos, ha sido necesario un gran presupuesto (450.000 millones de pesetas), un gran número de personas (en torno a 1.000 en todo el mundo), y el desarrollo de una gran tecnología y una elevada capacidad de automatización e informatización.

18.6 OTROS PROYECTOS DE SECUENCIACIÓN DEL GENOMA HUMANO

Como alternativas al PGH público, oficial, ha surgido en los últimos años una serie de proyectos, en cierto modo con un carácter marcadamente comercial, pero que han contribuido a adelantar la obtención de la secuencia del genoma humano, obviando en gran parte las etapas previas de mapeos genético y físico. Estos planteamientos han dado lugar a acalorada polémica ético-científica, ante el intento de apoderarse del lucrativo campo de la biotecnología, a medio camino entre lo científico y lo industrial, a modo de una batalla para hacerse con el dominio de la medicina y la biotecnología del futuro. El hecho de que incluso algunos investigadores estén fundando sus propias empresas, para explotar sus ideas antes de que lo hagan otros, ha creado una inquietud en la comunidad científica internacional y ha dado lugar a la intervención de los gobiernos financiadores en apoyo del acceso público y universal a todos los datos del genoma por parte de investigadores del sector académico o industrial farmacéutico.

Como ejemplo representativo, cabe citar la presencia de la empresa privada Celera Genomics (Rockville, EE.UU.) que, con contratos multimillonarios sin precedentes en la historia de la biología y la aplicación de un nuevo planteamiento de secuenciación, ha conseguido en sólo 3 años secuenciar el genoma humano, tras haber verificado previamente la viabilidad del método con la secuenciación del genoma de *Drosophila melanogaster*. La técnica empleada se puede resumir en 3 etapas:

1. **Secuenciación *shot-gun*** ("en escopetazo" o "en perdigonada"). Esta tecnología es muy diferente a la empleada por el PGH. Se basa en la escisión aleatoria de todo el genoma en pequeños fragmentos (unos 50 ó 70 millones), que se clonan para formar genotecas. Posteriormente, se procede a secuenciar los dos extremos de cada fragmento (en el mayor número posible de nucleótidos permitido por las técnicas de secuenciación disponibles). Esta ingente tarea, realizada de forma robotizada (se han requerido 300 secuenciadores automáticos funcionando día y noche), ha permitido secuenciar varias veces cada fragmento (4,6 de media, lo que supone una precisión inferior a la del PGH), hasta completar en sólo 3 años (abril de 2000) la secuencia del 99% de los fragmentos procedentes de un genoma de varón. El anuncio de estos resultados se hizo en un ambiente un tanto espectacular.
2. **Ensamblado de los fragmentos.** Consistió en un trabajo meramente informático, aunque de gran magnitud, encaminado a recomponer correctamente los fragmentos anteriores, a modo de puzzle gigante, hasta lograr en un solo mes el mapa completo. Se ha conseguido así, por primera vez en la historia (mayo de 2000), la secuencia de un genoma humano. Sin embargo, este resultado ha venido acompañado de fuertes críticas, por ser menos ordenado (aunque más rápido) que el método del PGH.
3. **Anotación (o identificación).** Es la fase necesaria para completar la secuencia de todos los fragmentos y tratar de localizar la posición de los genes en el genoma total. Es, precisamente, la más compleja (prevista su consecución para el año 2001) y a la vez la más importante. Equivale, en realidad, a la fase proteómica posterior a la genómica (pág. 217).

Estos dos proyectos, el público (PGH) y el privado (Celera), después de varios enfrentamientos verbales y críticas mutuas, pactaron finalmente un acuerdo, sin vencedores ni vencidos, acerca de los datos alcanzados por cada grupo. Como consecuencia, se ha elaborado un "borrador" común de la secuencia del genoma humano alcanzada bajo planteamientos distintos. Presentado el 26 de junio de 2000 (fecha ya histórica), a nivel internacional y con intervención de científicos y políticos, se pretende que este borrador sirva de partida para futuros estudios, presumiblemente realizados en adelante ya de forma conjunta.

Por otro lado, debe indicarse que algunos grupos, incluso antes de conocerse los resultados de la empresa pública y privada, han emprendido proyectos centrados en la secuenciación de los genes, exclusivamente, prescindiendo de las regiones no codificantes del genoma. Como ejemplo representativo, debe citarse la empresa Human Genomic Sciences que, desde 1992, ha conseguido secuenciar (hasta junio de 2000) más del 95% de los genes que se expresan. La técnica utilizada se basa en la transcripción inversa del mRNA de células específicas, para obtener su copia de cDNA, que luego es sometido a secuenciación. Su objetivo, intermedio entre lo científico (secuenciar sólo lo que es útil para la medicina) y lo comercial (fabricar medicamentos a partir de dichos genes), está adquiriendo una trayectoria dominada marcadamente por el aspecto económico, sin publicar los resultados de cara a la elaboración de patentes.

En cualquier caso, los avances conseguidos hasta ahora han supuesto, sin lugar a dudas, una revolución de gran impacto en la sociedad. De hecho, la aparición de los mapas y la secuenciación del genoma se han convertido en el centro de atención de los medios de comunicación. Pero ha de recordarse que la genómica "no es el final de un camino, sino el comienzo de todo". Una vez superados estos objetivos, se entrará en la fase proteómica o de deducción de la función de cada gen, su expresión y la aplicación de las proteínas formadas. Sólo así se podrán conocer las alteraciones proteicas responsables de muchas patologías y empezar a obtener beneficios del PGH, cifrados en campos tan diversos como la identificación, diagnóstico, terapia y prevención de muchas enfermedades, el conocimiento de la diversidad genética (polimorfismos del DNA), la elaboración de tratamientos individualizados, etc.

Queda igualmente pendiente un aspecto muy importante del genoma humano. Debe quedar claro que en ningún caso existirá una secuencia única para el genoma de todos los individuos de la especie humana, pues es bien conocida, desde hace muchos años, la existencia de una marcada diversidad o polimorfismo genético entre individuos de una misma especie. Este aspecto, que será analizado más adelante (tema 26), nos indica que los proyectos genoma, sobre todo en las regiones de DNA no codificante, deberán ser considerados bajo este nuevo prisma. Ello explica que los proyectos actualmente en curso ya estén tratando de abordar este aspecto con la secuenciación comparativa del DNA de individuos representativos de distintas razas, como un principio de distinción entre genomas polimórficos.

18.7 GENÓMICA COMPARATIVA

La interpretación del genoma humano se facilita por comparación de su secuencia y sus genes con los de otros organismos. En el momento actual se siguen desarrollando y se han iniciado o están planeados muchos otros proyectos genoma. En la lista se incluyen los de la levadura *Saccharomyces cerevisiae*, el gusano *Caenorhabditis elegans* o la mosca del vinagre *Drosophila melanogaster*. Entre los mamíferos, destaca el proyecto genoma del ratón, de gran interés aplicado en investigación y por la elevada homología con los genes humanos. El genoma de *Drosophila melanogaster*, recientemente secuenciado con la intervención (excepcional) de investigadores españoles, es uno de los modelos más utilizados para estudiar el control genético del desarrollo de organismos complejos, así como para el estudio de trastornos que afectan a humanos. Un 61% de los genes de *Drosophila* se han vinculado a enfermedades humanas, tales como el cáncer, el Parkinson, el Alzheimer y la diabetes. Igualmente, es importante el genoma de otras especies como modelos para el estudio de trastornos genéticos y sobre los fármacos a emplear. De ahí que los primeros genomas estudiados fueran los de microorganismos patógenos causantes de enfermedades humanas.

Muchos más genes de lo que se pensaba se encuentran en diversas especies animales. Con el tiempo su estudio, o genómica comparativa, permitirá avanzar en el conocimiento de procesos biológicos básicos: división celular, fertilidad, diferenciación sexual, desarrollo del cuerpo y órganos, metabolismo del envejecimiento, aprendizaje, susceptibilidad a fármacos, evolución, etc. Además, se aportará información sobre el mecanismo de la evolución; en este sentido, merece destacarse el descubrimiento de lo que podría ser el "genoma mínimo" necesario para mantener la vida de forma autónoma (300 del total de 470 genes de la bacteria *Mycoplasma genitalium*).

Genomas que se han secuenciado (mayo 2000) (ordenados por tamaño)			
Organismo	Características	Nº estimado de genes	Tamaño del genoma (Mb)
VIH-1 (virus de inmunodeficiencia humana)	virus		0,0097
DNA mitocondrial <i>Homo sapiens</i>	mitocondria	37	0,0165
Cromosoma 1 de <i>Leishmania major</i>	procariota		0,27
<i>Mycoplasma genitalium</i>	procariota	470	0,58
<i>Mycoplasma pneumoniae</i>	procariota	677	0,82
Cromosoma 2 de <i>Plasmodium falciparum</i>	procariota		1,0
<i>Chlamydia trachomatis</i>	procariota		1,05
<i>Rickettsia prowazekii</i>	procariota		1,1
<i>Treponema pallidum</i>	procariota	1.234	1,14
<i>Borrelia burgdorferi</i>	procariota	863	1,3
<i>Aquifex solicus</i>	procariota	1.508	1,5
<i>Methanococcus jannaschii</i>	arqueobacteria	1.692	1,66
<i>Helicobacter pylori</i>	procariota	1.590	1,67
<i>Halobacterium halobium</i>	procariota		1,7
<i>Methanobacterium thermoautotrophicum</i>	arqueobacteria	1.855	1,8
<i>Thermotoga maritima</i>	procariota		1,8
<i>Haemophilus influenzae</i>	procariota	1.740	1,83
<i>Pyrococcus furiosus</i>	procariota		2,1
<i>Archaeoglobus fulgidus</i>	arqueobacteria	2.471	2,2
<i>Pyrobaculum aerophilum</i>	procariota		2,2
<i>Deinococcus radiodurans</i>	procariota		3
<i>Synechocystis</i> sp.	procariota	3.168	3,57
<i>Clostridium acetobutylicum</i>	procariota		4,1
<i>Bacillus subtilis</i>	procariota	4.100	4,2
<i>Escherichia coli</i>	procariota	4.288	4,67
<i>Saccharomyces cerevisiae</i> (levadura de la cerveza y del pan)	eucariota	6.034	12,1
Brazo q del cromosoma 21, <i>Homo sapiens</i>	eucariota	225	33,6
Brazo q del cromosoma 22, <i>Homo sapiens</i>	eucariota	545	34,5
	(1º cromosoma humano secuenciado, 1999)		
<i>Caenorhabditis elegans</i> (gusano nematodo)	eucariota (1º genoma animal secuenciado, 1998)	17.000	100
<i>Drosophila melanogaster</i> (mosca de la fruta o del vinagre)	eucariota (el mayor genoma secuenciado)	13.600	180 (secuenciados 120, región eucromática)

Tema 19

Transcripción

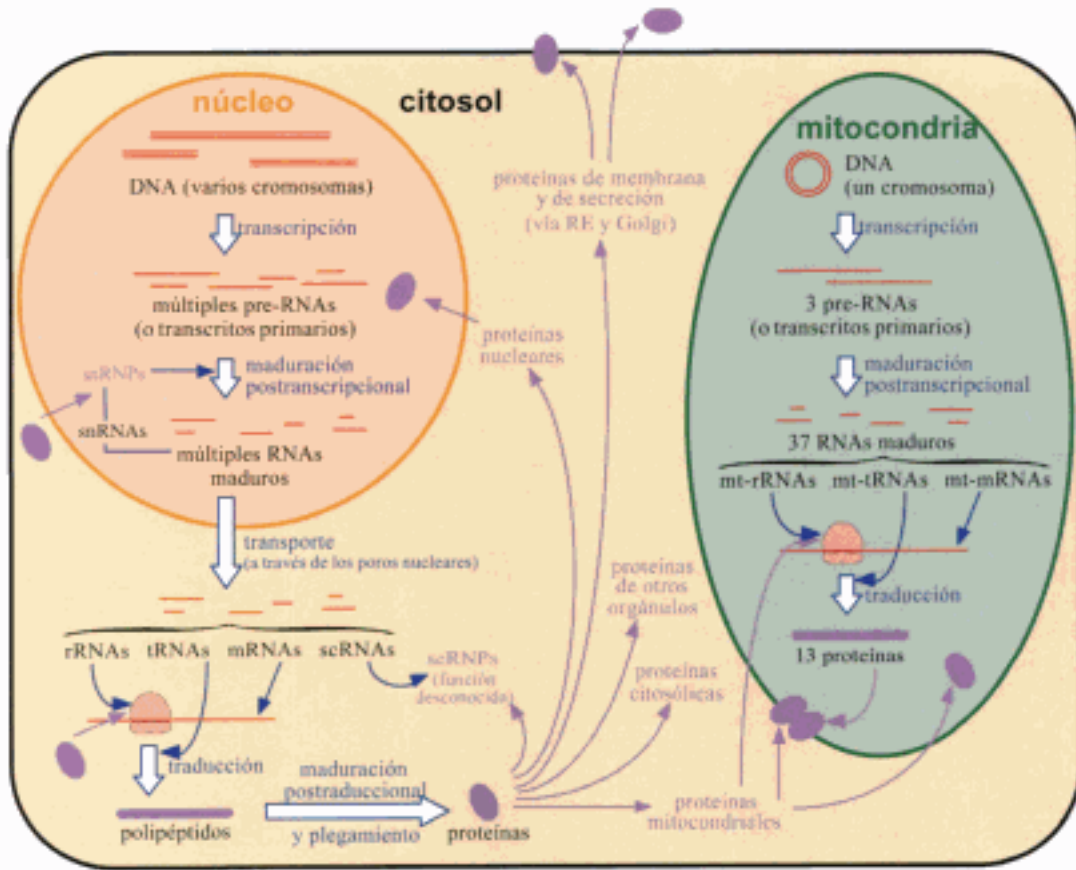
19.1	Introducción: conceptos generales	243
19.1.1	Transcripción y transcripción inversa	243
19.1.2	Concepto de gen	244
19.1.2.1	Definición clásica, no molecular	244
19.1.2.2	Definición molecular	244
19.1.2.3	Situaciones particulares	246
19.1.2.4	Nomenclatura	246
19.2	Enzimología de la transcripción: Mecanismo de la reacción RNA polimerasa	246
19.2.1	Terminología	248
19.2.2	Orientación de la secuencia	249
19.2.3	Carácter asimétrico	249
19.3	Transcripción en eucariotas: diferencias con procariotas	250
19.3.1	Relación con la condensación del DNA	250
19.3.2	Sólo se transcribe una pequeña parte del genoma	250
19.3.3	El proceso es mucho más complejo que en procariotas	250
19.3.4	Existen varias RNA polimerasas diferentes	250
19.3.4.1	Especificidad respecto al tipo de gen transcrito y RNA formado	251
19.3.4.2	Localización y cantidad	251
19.3.4.3	Susceptibilidad a inhibidores	251
19.3.5	Las RNA polimerasas están formadas por un gran número de subunidades	251
19.4	Etapas en el proceso de transcripción	252
19.4.1	Iniciación	253
19.4.1.1	Formación del complejo de iniciación	253
19.4.1.2	Desnaturalización del promotor	253
19.4.1.3	Formación de los primeros enlaces fosfodiéster	254
19.4.1.4	Liberación del promotor: transición de iniciación a elongación	254
19.4.2	Elongación o alargamiento del RNA	254
19.4.3	Terminación	255
19.4.3.1	Terminación en procariotas	255
19.4.3.2	Terminación en eucariotas	256
19.5	Transcripción del genoma mitocondrial	257
19.6	Inhibidores de la transcripción	258
19.6.1	Antibióticos de tipo I	258
19.6.2	Antibióticos de tipo II	258
19.6.3	Antibióticos de tipo III	259

19.1 INTRODUCCIÓN: CONCEPTOS GENERALES

19.1.1 Transcripción y transcripción inversa

La transcripción es el proceso encargado de la síntesis de una molécula de RNA a partir de la información genética contenida en la región codificante de un DNA. Es decir, de dar lugar a una "copia" de RNA (con secuencia no idéntica, sino complementaria y antiparalela) a partir de una secuencia "molde" en una de las hebras del DNA. Este proceso constituye el segundo paso del esquema clásico de transmisión de la información genética (pág. 1).

Se trata de un proceso enzimático catalizado en todos los organismos por una enzima RNA polimerasa (polimerasa de RNA dependiente de DNA, o transcriptasa, pág. 147). Como excepción, muchos virus con RNA como material genético (por ejemplo, los retrovirus) utilizan su propio RNA como molde para la síntesis de un DNA complementario (cDNA), mediante una polimerasa de DNA dependiente de RNA (transcriptasa inversa, págs. 147 y 195). Por esta razón, el esquema general del flujo de información genética se considera actualmente que es bidireccional en su primera parte (pág. 2). La presencia del OH 2' en la ribosa, que hace al RNA más reactivo (pág. 35), puede explicar que el papel del RNA como material genético (probable en todos los organismos al inicio de la evolución) se haya desplazado al DNA, una molécula más estable.



19.1.2 Concepto de gen

19.1.2.1 Definición clásica, no molecular

En términos genéticos clásicos, el gen se define como la unidad elemental de la herencia, la región física y funcional que controla una característica hereditaria concreta, la portadora de la información genética de una generación a la siguiente, la que gobierna, en definitiva, las características de un rasgo particular.

19.1.2.2 Definición molecular

En términos moleculares, puede distinguirse entre la visión original y el concepto actual.

En una fase previa, la definición de gen se basó en planteamientos muy simples, relacionados con los avances producidos con el tiempo:

Hipótesis un gen—una enzima. Basada en el papel del gen como responsable de la síntesis de enzimas, primeras moléculas con funciones catalíticas en las células.

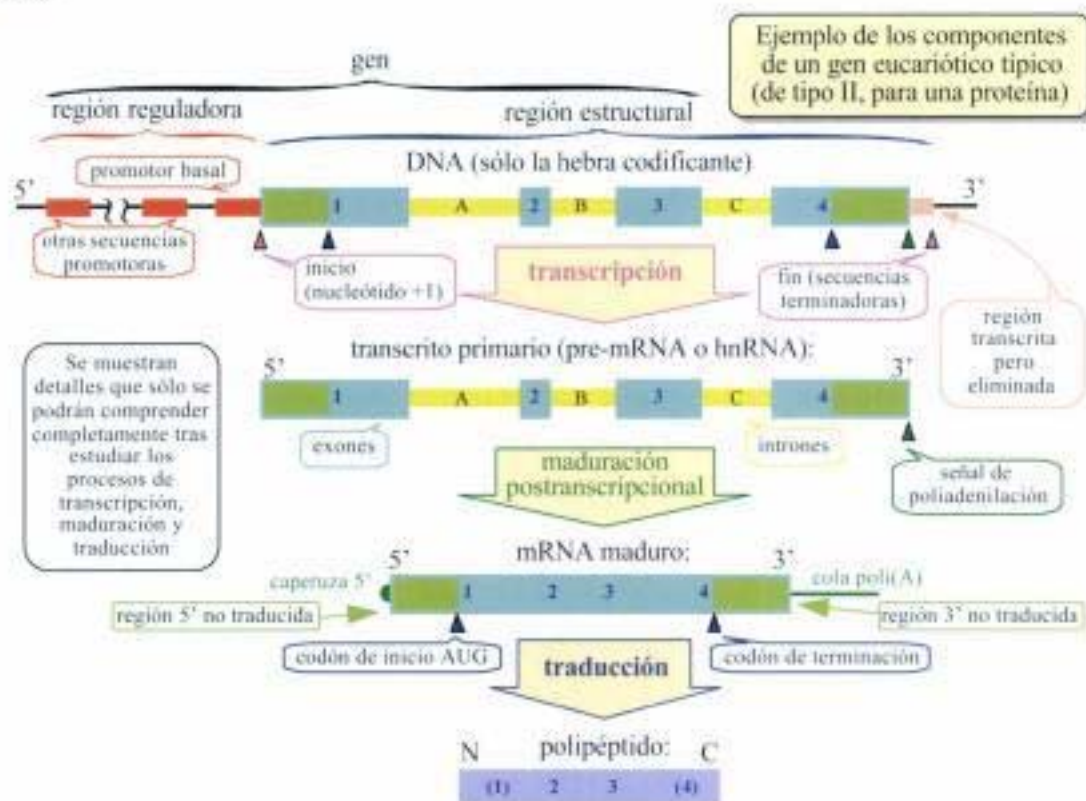
Hipótesis un gen—una proteína. Basada en la noción original de que todas las enzimas son proteínas (hoy se sabe que algunas moléculas no proteicas, como ribozimas, también tienen funciones catalíticas).

Hipótesis un gen—un polipéptido. Basada en que muchas proteínas están formadas por varios polipéptidos, cada uno codificado por un gen diferente.

Como consecuencia, la definición básica, la considerada normalmente como válida y aceptada, es:

Un gen es aquella región del genoma que contiene la información necesaria para sintetizar una molécula de polipéptido.

En el momento actual, este concepto debe ser matizado y ampliado para dar entrada a dos nuevas características de un gen cualquiera: existencia de dos tipos de secuencias con funciones diferentes, una estructural o realmente codificadora y otra reguladora de la expresión, y presencia en la región estructural de regiones codificantes y no codificantes.



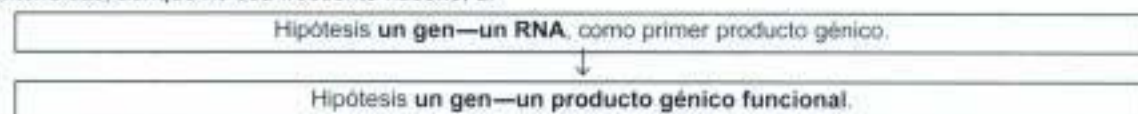
La **región estructural** determina la expresión real del gen. Comprende dos tipos de secuencias, en función de su capacidad de expresión: **intrones**, o regiones no codificantes presentes en el interior del gen (pág. 276), y **exones**, que incluyen tanto las secuencias codificantes como las no codificantes de ambos extremos del gen. El conjunto de exones e intrones de la región estructural se transcribe para dar lugar a un RNA llamado **precursor** o **transcrito primario**. Éste requiere un proceso adicional, posterior a la transcripción, para dar las moléculas funcionales de RNA (tema 21). La mayor parte de los transcritos formados en procariotas y eucariotas sufre dicho proceso, llamado **maduración postranscripcional**, siendo excepción los mRNAs procarióticos, que se sintetizan directamente en su forma funcional. Como parte de ese proceso se pierden los intrones, y los exones se unen linealmente, hasta constituir el RNA maduro o funcional (pág. 280).

La **región reguladora**, sin función codificante, está situada normalmente corriente arriba (es decir, hacia 5', pág. 249) de la región estructural. Contiene distintas **regiones promotoras**, encargadas de interactuar con los factores de transcripción proteicos para regular positiva o negativamente el inicio de la transcripción (pág. 264).

Lo importante para la definición molecular del gen es la formación a partir del DNA de dos tipos de productos génicos: RNAs o proteínas funcionales. Para la formación de los primeros se requiere la transcripción y, en su caso, la maduración postranscripcional. Sólo uno de los tipos de RNA, el mensajero, sirve como punto de partida para sintetizar el segundo producto génico, la proteína funcional. Ello ha requerido, además de la transcripción y maduración postranscripcional, la traducción o síntesis del polipéptido y, en la mayoría de los casos, el plegamiento proteico, necesario para su tráfico al lugar de acción y para el ejercicio de su función.

Como consecuencia, la definición molecular actual se ha ampliado: **un gen es el conjunto de secuencias de DNA de todo tipo, estructurales (intrones y exones) y reguladoras, necesarias para codificar un producto génico, sea éste un RNA maduro de cualquier tipo o una proteína funcional.**

Obsérvese que ambas definiciones moleculares están basadas en la hipótesis *un gen - un polipéptido*, pueden ser extendidas, aunque no sea frecuente hacerlo, a:



19.1.2.3 Situaciones particulares

Algunos genes poseen características que requieren forzar las definiciones establecidas; en esos casos, se debe acudir a la interpretación más funcional del concepto moderno de gen, más que a los detalles de la definición. Se indican algunos ejemplos.

- Dos genes pueden ser solapantes, es decir, compartir una misma región de DNA (pág. 249), bien porque cada uno está codificado en una hebra distinta (y, por tanto, sus secuencias son completamente diferentes, pág. 44) o bien incluso en la misma hebra (lo cual no implica que las proteínas codificadas compartan secuencia aminoacídica, debido a la existencia de tres marcos de lectura, pág. 288).
- Algunos genes dan lugar a varios productos (pág. 276). Un ejemplo es el gen de los rRNAs, que se transcribe en un RNA precursor que por maduración se fragmenta dando varios rRNAs funcionales (pág. 283). Otro caso está en el DNA mitocondrial, cuyos 3 transcritos primarios maduran para dar cada uno varios RNAs (pág. 284). Finalmente, muchos genes procarióticos se transcriben a un mRNA único que sin embargo da lugar por traducción a varias proteínas (mRNAs policistronicos, pág. 301).
- En otros casos, una misma secuencia de DNA puede dar lugar alternativamente a dos productos génicos, a través de variantes de su expresión génica; por ejemplo, por un diferente procesamiento postranscripcional o postraduccional (temas 21 y 24).
- Igualmente, en numerosos casos la proteína está formada por varias subunidades; en este caso, no existe un gen para la proteína, sino para cada uno de esos polipéptidos (por ejemplo, no se habla del gen de hemoglobina, sino de los genes de la α -globina y la β -globina).
- En casos excepcionales, como las inmunoglobulinas, un gen único sufre reorganizaciones internas en su secuencia antes de ser transcrito, de forma que da lugar directamente a una variedad de polipéptidos (los distintos anticuerpos con especificidad para antígenos diversos).

19.1.2.4 Nomenclatura

Los genes se nombran generalmente con tres letras minúsculas y en cursiva (por ejemplo, gen *ras*). La proteína codificada por él recibe a veces el mismo nombre pero comenzando con una letra mayúscula, y sin cursiva (proteína *Ras*). En ocasiones se emplean las tres letras en mayúscula (por ejemplo, gen *APC*).

19.2 ENZIMOLOGÍA DE LA TRANSCRIPCIÓN: MECANISMO DE LA REACCIÓN RNA POLIMERASA

Como ya se ha indicado, todos los organismos, procariotas y eucariotas, sintetizan el RNA de acuerdo con una reacción catalizada por la RNA polimerasa.

Previamente, por su significado histórico en relación con la síntesis de RNA, debe citarse la enzima **polinucleótido fosforilasa** bacteriana (descrita por Marianne Grunberg-Manago y Severo Ochoa en 1955, y llamada así por su analogía con la fosforilasa hepática). Esta enzima requiere un cebador y Mg^{+2} , pero al no utilizar un DNA molde conduce a una polimerización lineal al azar, no predeterminable. La polimerización se produce por unión de muchos 5'-NDPs mediante enlaces 3'-5'-fosfodiéster con la correspondiente liberación de P_i , para formar una molécula semejante al RNA natural (un polirribonucleótido):



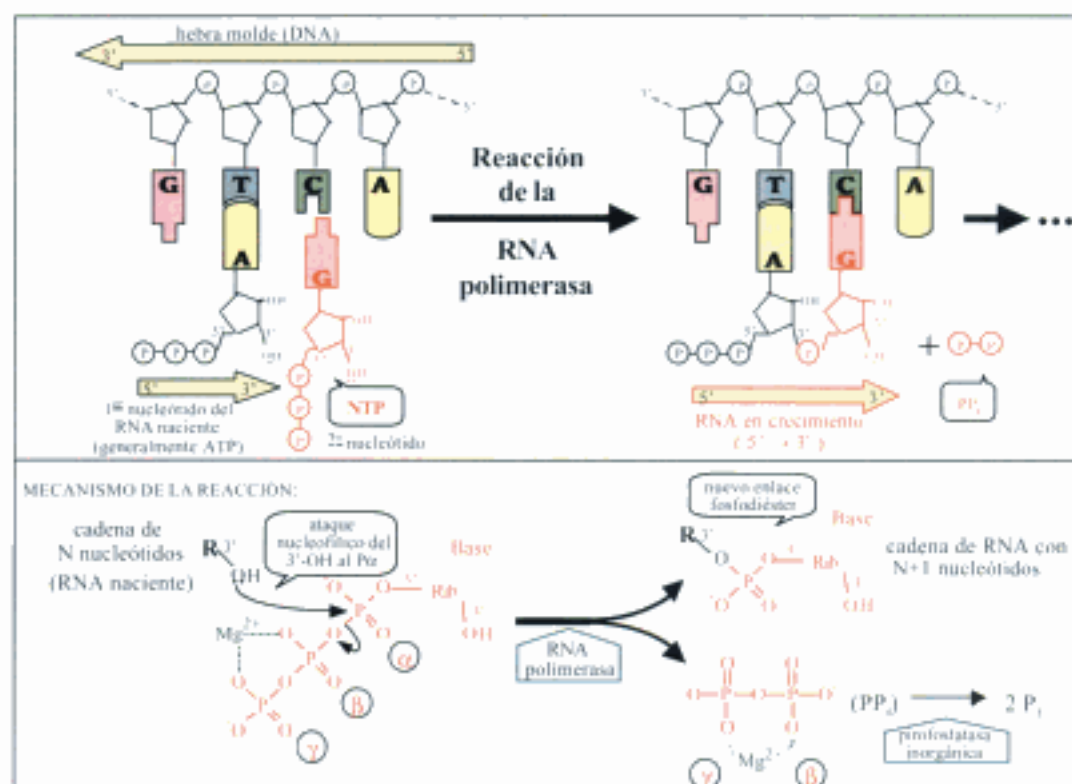
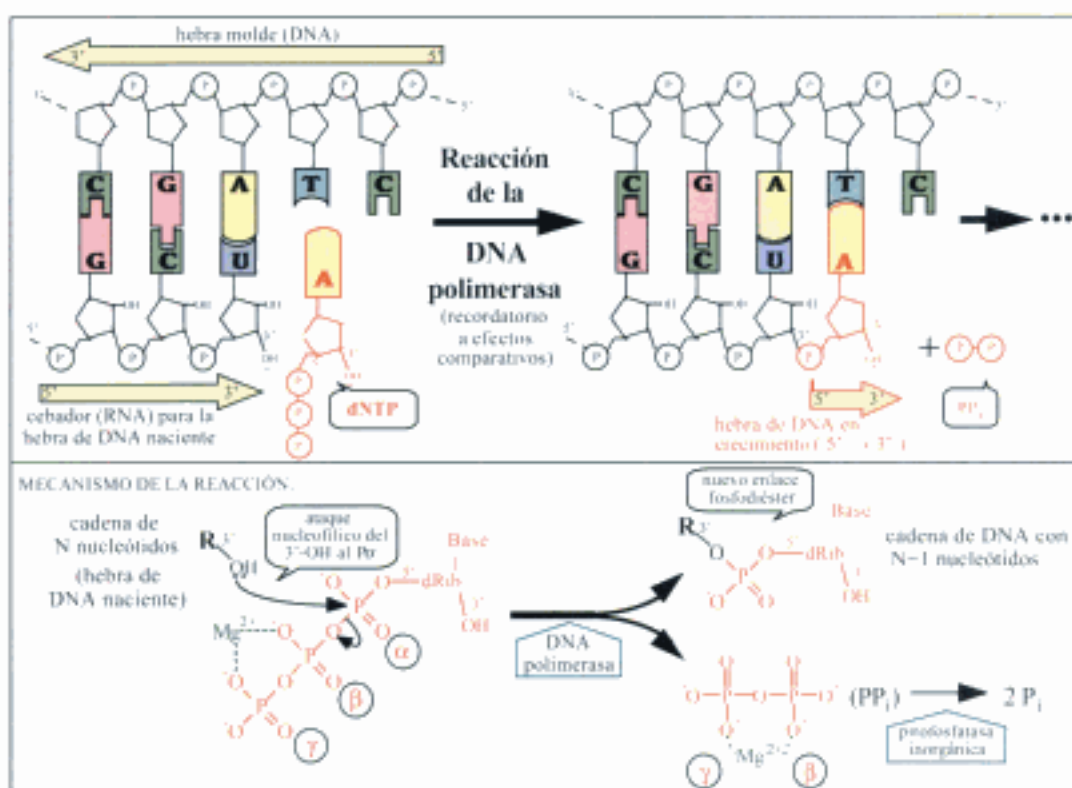
La incorporación de cada nucleótido depende de su concentración relativa (respecto a los otros tres) en la mezcla de reacción. Además, la posición de equilibrio de la reacción favorece más la fosforólisis que la polimerización. No se conoce que las células utilicen esta enzima; se cree que puede servir más bien como mecanismo de degradación que de síntesis de RNA.

El mecanismo de la reacción fisiológica, catalizada por la transcriptasa, es básicamente idéntico al de la DNA polimerasa (pág. 146). La **reacción global** es la siguiente:



El descenso de energía libre de la reacción, para cada nucleótido añadido, es $\Delta G^\circ = -2.6 \text{ kcal/mol} = -9 \text{ kJ/mol}$. Es decir, es ligeramente exergónica y espontánea. Al igual que en la DNA polimerasa, la reacción está asistida termodinámicamente por la ruptura del enlace fosfoanhidrido del NTP y la hidrólisis del PP_i por la pirofosfatasa. Se asegura así que la reacción transcurra irreversiblemente en la dirección de síntesis.

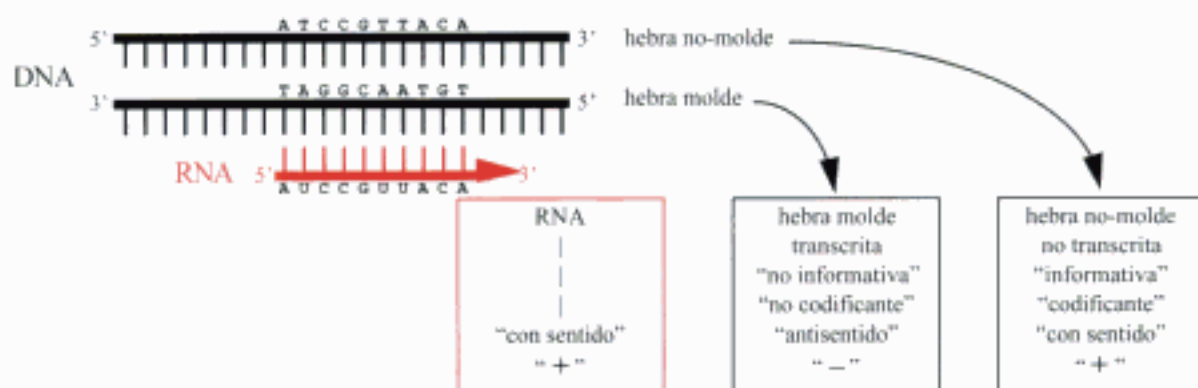
La adición de nucleótidos a la hebra de RNA nueva y las características de la reacción se exponen comparativamente con las de la replicación:



	Reacción de la DNA polimerasa		Reacción de la RNA polimerasa
1	Requiere una estructura desenrollada.	=	Requiere también una estructura desenrollada.
2	La enzima toma instrucciones de una hebra de DNA (molde).	=	También necesita como molde una hebra de DNA. No sirve el RNA.
3	Las dos hebras del DNA actúan como molde para la síntesis del nuevo DNA, simultáneamente. Cada DNA dúplex formado conserva una de las hebras del DNA original.	≠	Las dos hebras del DNA pueden actuar como molde del RNA, pero no simultáneamente. Tras realizarse la copia de una sola hebra (la molde), se conserva íntegro el DNA de doble hebra.
4	Utiliza cuatro dNTPs como sustrato: dATP, dGTP, dTTP y dCTP.	≠	Utiliza cuatro NTPs como sustrato: ATP, GTP, UTP y CTP.
5	No es autoiniciadora. Para iniciar la síntesis requiere RNA como cebador, que se aparea por complementariedad de bases y antiparalelamente sobre el DNA molde.	≠	Sí es autoiniciadora, no requiere cebador. Se inicia en región promotora, con un NTP (generalmente ATP) apareado al molde. El NTP mantiene su tri-P-5' y su 3'-OH libre.
6	La síntesis comienza por el extremo 3'-OH del cebador. La enzima asociada al DNA molde hace que este 3'-OH libre ataque nucleofílicamente al α -P del primer dNTP, ya apareado. Se forma el 1 ^{er} enlace fosfodiéster, se rompe el fosfoanhidrido α - β y se libera PP _i . Su hidrólisis favorece la reacción de síntesis.	=	La síntesis comienza por el 3'-OH del 1 ^{er} NTP apareado. La enzima asociada al DNA molde hace que este 3'-OH ataque nucleofílicamente al α -P del 2 ^o NTP, ya apareado. Se forma el 1 ^{er} enlace fosfodiéster, se rompe el fosfoanhidrido α - β y se libera PP _i , cuya hidrólisis favorece la reacción de síntesis.
7	Requiere Mg ⁺² como cofactor enzimático para que la actividad sea óptima.	=	Necesita un ion metálico divalente. Sirven Mn ⁺² y Mg ⁺² , pero este último actúa <i>in vivo</i> .
8	Se producen apareamientos definitivos A=T y C≡G entre la hebra de DNA nuevo y la de DNA molde.	≈	Se producen apareamientos temporales A=U y C≡G entre RNA y DNA molde.
9	Es reiterativa (elongación). Se repite la adición de dNMPs por ataques nucleofílicos del 3'-OH del DNA nuevo sobre el α -P-5' de cada dNTP apareado. La síntesis es en dirección 5'→3' de la nueva cadena de DNA, es decir, en dirección 3'→5' del DNA molde.	=	Idem. El proceso se repite. La adición de NMPs es por ataques nucleofílicos del 3'-OH del RNA nuevo sobre el α -P en 5' de cada NTP apareado. Igualmente, la síntesis es en dirección 5'→3' del RNA naciente, es decir, en dirección 3'→5' del DNA molde.

19.2.1 Terminología

Para una descripción didáctica de la transcripción es aconsejable trazar en horizontal el DNA de doble hebra que se va a transcribir; se considera a la hebra superior (extremo 5'-P a la izquierda) como **hebra no-molde** y a la inferior, antiparalela (extremo 5'-P a la derecha), como **hebra molde**. Volviendo ésta, pasa a ser la superior en su apareamiento con el RNA sintetizado, que se sitúa, antiparalelamente (extremo 5'-PPP a la izquierda y 3'-OH a la derecha), debajo de esa hebra molde. La **síntesis del RNA**, que se marca a menudo con línea (inicio) y flecha (sentido), **ocurre en sentido 5'→3'** del RNA naciente. Esto equivale a decir que la hebra molde de DNA se transcribe (se "lee") en dirección 3'→5'.



La secuencia del RNA nuevo es, por tanto, idéntica a la de la hebra de **DNA no-molde**, salvo por la presencia de U en lugar de T. Por ello, esta hebra se llama **codificante** o **informativa** (corresponde a la secuencia de bases del RNA y originará, por traducción, la secuencia de aminoácidos de la proteína). También se conoce como hebra **no transcrita**, hebra **(con) sentido** y hebra **positiva (+)**. La secuencia de esta hebra es la que normalmente se documenta al dar la secuencia de un gen. En consonancia con lo anterior, la hebra de **DNA molde** se llama hebra **no codificante**, **no informativa**, **transcrita**, **antisentido** o **negativa (-)**.

19.3 TRANSCRIPCIÓN EN EUCARIOTAS: DIFERENCIAS CON PROCARIOTAS

19.3.1 Relación con la condensación del DNA

Al igual que en la replicación eucariótica (pág. 151), un problema especialmente importante es el nivel de condensación del DNA al inicio del proceso de transcripción. En los cromosomas metafásicos, donde la condensación es máxima, está totalmente impedida la transcripción. Ésta tiene lugar durante la interfase, pues sólo ocurre sobre la conformación de fibras de 10 nm, de mínima condensación (pág. 85) y, por tanto, mayor accesibilidad a las RNA polimerasas y otras proteínas necesarias para el proceso. La mayor descondensación nunca se da simultáneamente a todo lo largo del cromosoma, lo que explica la correlación observada experimentalmente entre actividad transcripcional y posición de las regiones descondensadas (eucromatina, pág. 87).

19.3.2 Sólo se transcribe una pequeña parte del genoma

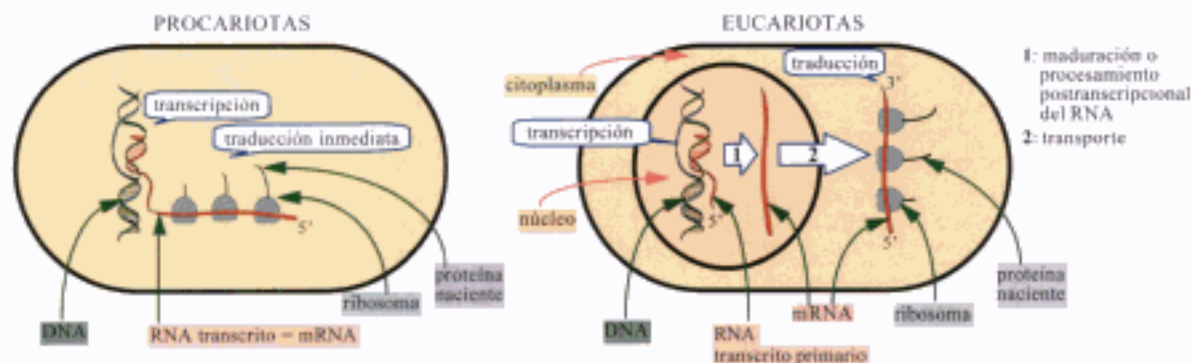
Como ya se indicó, una gran mayoría del DNA genómico de eucariotas nunca se transcribe (DNA no codificante, en buena parte repetitivo, pág. 108). Asociadas a este carácter no funcional o simplemente no transcribible (pág. 107) aparecen copias defectuosas de genes (pseudogenes y fragmentos de genes, pág. 112), estructuras cromosómicas que no se transcriben, etc. Sólo se transcribe el 35% restante, correspondiente a los exones de las regiones estructurales de los genes, sean de secuencia única o repetida, y a sus intrones (aunque éstos, estrictamente, son DNA no codificante). Además, dentro del DNA que se puede transcribir, células diferentes transcriben distintas regiones de DNA, según sus necesidades y gracias al complejo proceso que constituye el control de la expresión génica (tema 20).

En este contexto parece adecuado insistir en que no es válida actualmente la definición de gen como "secuencia de DNA que se transcribe", porque excluye la región del genoma que controla la transcripción (pág. 244).

19.3.3 El proceso es mucho más complejo que en procariotas

En **procariotas**, la molécula de mRNA resultante de la transcripción (transcrito primario) es ya funcional, no precisa sufrir el proceso de maduración postranscripcional y se utiliza inmediatamente como molde para la traducción, en el mismo compartimento subcelular (puesto que no hay membrana nuclear). Existe por ello una asociación *temporal y espacial* íntima entre transcripción y traducción; de hecho, un mRNA puede empezar a traducirse antes de haberse completado su síntesis por transcripción. Por el contrario, la mayoría de los tRNAs y rRNAs deben sufrir maduración, análoga a la de eucariotas.

En **eucariotas**, el transcrito primario, o RNA resultante de la transcripción, experimenta en el núcleo la *maduración o procesamiento postranscripcional*. Los RNAs maduros se transportan luego al citosol para participar en la traducción. Existe, por tanto, una *separación espacial y temporal* entre transcripción y traducción, que supone una regulación más compleja de la expresión génica, contribuyendo a la riqueza en variedad y función propia de las células eucariotas. Este control se lleva a cabo en gran medida por regulación de la actividad de la RNAPol-II. El hecho de que esta enzima no se una específicamente por sí misma a los promotores eucarióticos es otro signo de la complejidad respecto a procariotas.



19.3.4 Existen varias RNA polimerasas diferentes

En **procariotas y orgánulos** subcelulares (mitocondrias y cloroplastos de eucariotas) se utiliza una sola RNA polimerasa para sintetizar los tres tipos de RNA celular (mRNA, tRNA y rRNA).

En el núcleo de **eucariotas** existen tres RNA polimerasas diferentes (I, II y III), que sintetizan cada una distintos tipos de RNA. Sus características diferenciales les confieren un mayor grado de especialización en cuanto a su labor de síntesis de RNA. Además, emplean distintos factores de transcripción que se unen a promotores situados en distinta posición (tema 20).

RNA polimerasas eucarióticas					
	Transcrito primario que sintetiza	Producto génico final	Localización	Actividad enzimática	Efecto de la α -amanitina
RNApol-I	pre-rRNA grande	rRNAs de 28S, 18S y 5.8S	Nucleolo	50-70%	Ninguno
RNApol-II	hnRNAs (pre-mRNAs) (*) y mayoría de snRNAs	mRNAs	Nucleoplasma	20-40%	Fuerte inhibición
RNApol-III	pre-tRNAs, pre-rRNA pequeño (*) y varios RNAs pequeños	tRNAs rRNA de 5S	Nucleoplasma	~10 %	Débil inhibición
RNApol	(*) rRNAs, mRNAs y tRNAs mitocondriales o cloroplásticos		Mitocondrias y cloroplastos		Ninguno

(*) no se conoce si se sintetizan en forma activa directamente o como pre-RNAs que necesitan maduración

19.3.4.1 Especificidad respecto al tipo de gen transcrito y RNA formado

Cada RNApol transcribe un tipo de gen y da lugar a transcritos primarios que originarán por maduración posttranscripcional distintos tipos de RNA maduros. Por ahora se comenta sólo el primero de estos aspectos.

- La RNApol-I transcribe el gen del pre-rRNA "grande" (45S), precursor único de 3 de los 4 tipos de rRNAs componentes del ribosoma citoplasmático (págs. 70 y 80).
- La RNApol-II transcribe todos los genes de proteínas, sintetizando los RNAs nucleares heterogéneos (hnRNAs), que son los precursores de los mRNAs (por ello, el hnRNA está constituido por moléculas de tamaños muy variados, de lo que procede su nombre). La misma RNApol-II transcribe también la mayoría de genes de los RNAs nucleares de tamaño pequeño (snRNA). Puesto que RNApol-II transcribe la gran mayoría de genes, los que codifican proteínas, esta enzima se considera representativa de las polimerasas eucarióticas.
- La RNApol-III, por su parte, transcribe los genes de todos los tRNAs, sintetizando sus pre-tRNAs respectivos, así como el gen del pre-rRNA "pequeño", que madurará para dar lugar al cuarto tipo de rRNA (5S), y los de otros RNAs pequeños.

19.3.4.2 Localización y cantidad

De acuerdo con su especificidad, las RNApol se localizan en el nucleoplasma (RNApol-II y III) o en el nucleolo (RNApol-I). Éste es una región densa del núcleo, de tinción más intensa, rica en RNA y proteínas, no rodeada por membrana alguna ni observable durante la mitosis. Surge de la propia transcripción de los genes de rRNA, proceso muy activo para proporcionar la elevada cantidad de ribosomas que necesita la célula.

La cantidad de cada tipo de RNApol (medida como fracción de la actividad enzimática total) refleja la demanda de sus productos génicos respectivos: es máxima para RNApol-I, intermedia la de RNApol-II y casi residual en RNApol-III.

19.3.4.3 Susceptibilidad a inhibidores

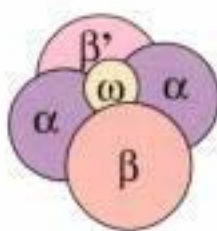
Es la diferencia más significativa, al permitir la distinción experimental de las 3 polimerasas nucleares por su respuesta a la α -amanitina (toxina de *Amanita phalloides*), un octapéptido bicíclico con varios aminoácidos infrecuentes (pág. 258). En células animales este compuesto no afecta a la RNApol-I, inhibe fuertemente (a baja concentración, 10^{-6} M) a la RNApol-II, y débilmente a la RNApol-III (a mayor concentración, entre 10^{-5} y 10^{-6} M; aunque de forma variable según la especie).

En cuanto a la RNApol mitocondrial, no se ve afectada por la α -amanitina, pero sí por la **rifampicina**, que no afecta a las polimerasas nucleares. En la respuesta a ambos inhibidores la polimerasa mitocondrial se comporta igual que la RNApol de células procarióticas (recuérdese el probable origen endosimbiótico de las mitocondrias, pág. 8).

19.3.5 Las RNA polimerasas están formadas por un gran número de subunidades

Esta es una característica común, aunque con ciertas diferencias entre procariotas y eucariotas.

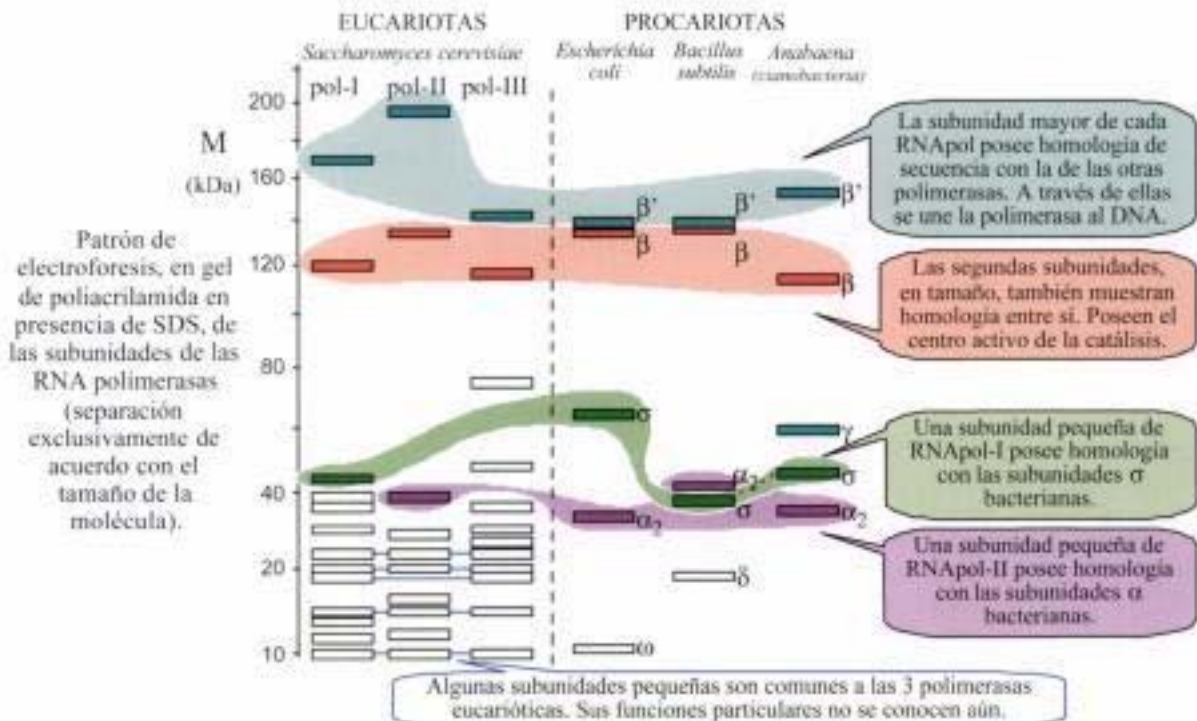
La RNApol de **procariotas** (*E. coli*), la mejor estudiada, está formada por 4 (posiblemente 5) polipéptidos, asociados en la llamada enzima "minima" o "núcleo". Para iniciar la transcripción, la enzima mínima requiere la unión de otro polipéptido independiente, la subunidad o factor σ , que participa reconociendo las secuencias promotoras en el DNA, o sitio de inicio de la transcripción. Una vez iniciada la transcripción, la subunidad σ se vuelve a disociar de la enzima mínima, que continúa elongando el RNA.



Estructura de una RNA polimerasa procariótica (*E. coli*)

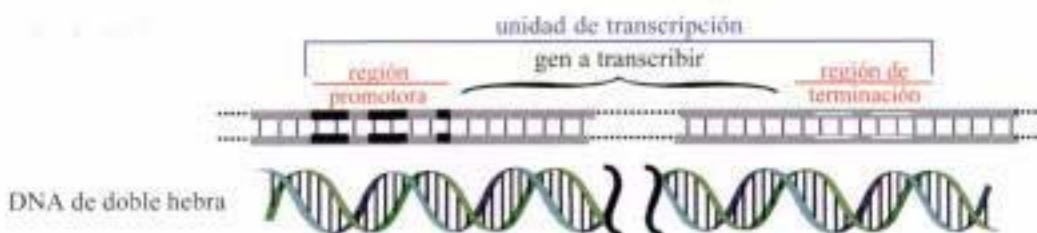
- | | | | |
|--------------|---|--------------------------------------|--|
| α (36.5 kDa) | = ensambla la enzima mínima y se une a proteínas reguladoras | enzima "núcleo" o "mínima" (390 kDa) | |
| β (151 kDa) | = polimerasa (sitio catalítico) | | |
| β' (155 kDa) | = se une al DNA molde | | |
| ω (11 kDa) | = (posible, no confirmado) | | |
| σ (70 kDa) | = especificidad por secuencias de inicio: reconoce al promotor e inicia la síntesis | | |

Las RNAPol de **eucariotas** son mucho más complejas. Aparecen como grandes agregados de masa cercana a 600 kDa, formados por entre 10 y 14 subunidades. La subunidad mayor de la RNAPol-II contiene un dominio C-terminal (llamado **CTD**, de *C-terminal domain*) con múltiples repeticiones (50 en mamíferos) de la secuencia Tyr-Ser-Pro-Thr-Ser-Pro-Thr. Este dominio puede ser fosforilado en Ser y Thr, adquiriendo una elevada carga negativa, lo cual está relacionado con su función en la transición entre las fases de iniciación y elongación (pág. 254).



19.4 ETAPAS EN EL PROCESO DE TRANSCRIPCIÓN

Para que tenga lugar la transcripción es necesaria la presencia en la hebra molde de la llamada **unidad de transcripción**, que incluye la secuencia del DNA que se ha de transcribir más las dos secuencias consenso que la flanquean (llamadas **promotor** y **terminador**). Es impropia la definición de unidad de transcripción como "región del DNA que se extiende entre el promotor y el terminador". Es igualmente incorrecto considerar al promotor como la "región del DNA que se transcribe". Se reserva el término **operón** para la unidad de transcripción en procariotas.



Al igual que la replicación, la transcripción en eucariotas transcurre en las tres fases clásicas de iniciación, elongación y terminación, con notables diferencias respecto a procariotas en cuanto al reconocimiento de las secuencias promotora y terminadora.

19.4.1 Iniciación

Es, con mucho, la etapa más compleja de la transcripción (y, por ello, en la que tiene lugar la mayor parte de la regulación, tema 20). Requiere la separación de las hebras del DNA en un pequeño tramo cercano a la **secuencia promotora**, para permitir que la RNA polimerasa sintetice un fragmento corto de RNA (de unos 10 nucleótidos) por apareamiento con la hebra molde. La región de DNA que sufre este desenrollamiento parcial se llama **burbuja de transcripción**; ésta se forma en el inicio y posteriormente, durante la elongación, se desplaza a lo largo del DNA junto con la RNAPol. Para la iniciación también se requiere la unión al DNA de proteínas llamadas factores de inicio de la transcripción o, comúnmente, **factores de transcripción (TF)**. Tanto éstos como los promotores se estudiarán en detalle más adelante (tema 20).

En **procariotas** y en las **mitocondrias** es frecuente que un mismo promotor inicie la **cotranscripción** de varios genes contiguos. Por el contrario, en el genoma **nuclear de eucariotas**, a partir de cada promotor sólo se transcribe un gen. Distintos genes, en especial en función de la RNAPol, utilizan distintos tipos de secuencias promotoras, a las que se unen distintos factores de transcripción. El proceso de iniciación que se describe como típico de eucariotas utiliza promotores que tienen la secuencia TATA, pero no la Inr (llamados por ello de tipo TATA+ Inr-), cuyas características se describen posteriormente (pág. 265). La secuencia TATA es reconocida por la RNAPol indirectamente, previa interacción con los TFs. Es importante resaltar que tanto en procariotas como en eucariotas, la iniciación requiere la presencia de la RNAPol completa u **holoenzima** (es decir, RNAPol bacteriana con la subunidad σ , pág. 252, y RNAPol eucarióticas con todas sus subunidades y con los factores de transcripción).

La iniciación puede subdividirse en tres subetapas, que sólo se describen para los genes transcritos por la RNAPol-II (genes de clase II). Los genes de clase I y III se transcriben por mecanismos similares, salvo por la identidad y peculiaridades de sus RNAPol, promotores y TFs.

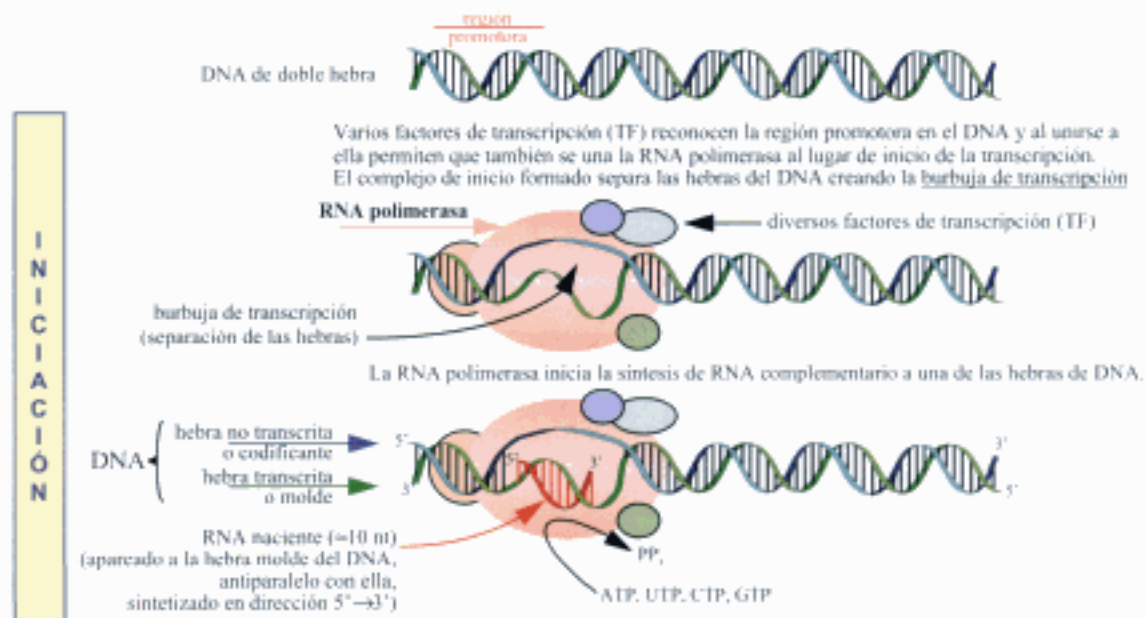
19.4.1.1 Formación del complejo de iniciación

Este complejo plurimolecular es **imprescindible para que la RNAPol pueda actuar**. Supone la unión de varios TFs a las regiones promotoras del DNA y la incorporación de la enzima, que participa en la forma con dominio CTD no fosforilado (pág. 252). Se han propuesto dos modelos para la formación de este complejo, con el mismo resultado; sus detalles se describen posteriormente, una vez estudiados los promotores y los TFs (pág. 267).

- El **modelo del holoenzima**, en el que la RNAPol-II se une previamente a los TFs y luego el conjunto se une al promotor.
- El **modelo escalonado**, en el que los TFs se van asociando a la región promotora uno a continuación del otro y en un cierto orden. Es el modelo actualmente más aceptado.

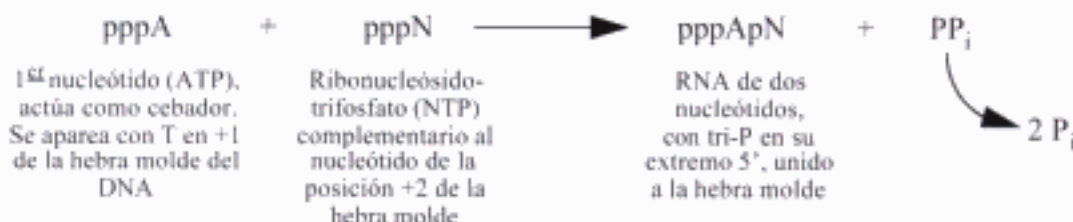
19.4.1.2 Desnaturalización del promotor

El complejo de iniciación produce en la región de inicio de la transcripción el desenrollamiento y separación de las dos hebras del DNA (desnaturalización local), esencial para la exposición de las bases a la RNAPol. Esta zona de hebras separadas, **burbuja de transcripción** o **complejo abierto**, es el lugar donde va a tener lugar la síntesis de la cadena de RNA. Esta etapa requiere, al igual que el inicio de la replicación (pág. 151), una actividad **DNA-helicasa**, presente ahora en dos de los factores de transcripción, TFIIIF y TFIIH, con un gasto energético (hidrólisis del ATP). Otro factor, el TFIIIE, estabiliza la región desnaturalizada del promotor (papel en cierto modo similar al ejercido por la RPA en la replicación, pág. 151).



19.4.1.3 Formación de los primeros enlaces fosfodiéster

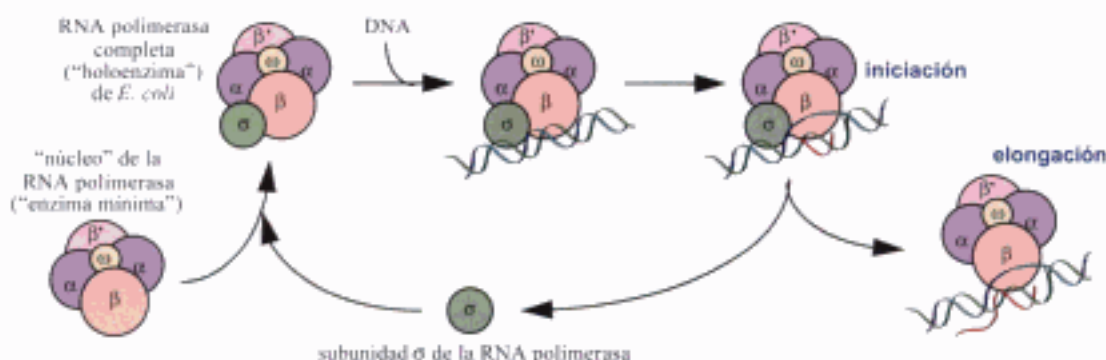
El enlace fosfodiéster inicial se forma entre dos ribonucleótidos, a partir de sus NTPs en forma libre. Uno de ellos, generalmente con base purínica (concretamente ATP), se aparea con el nucleótido **+1** en la hebra molde de DNA. El otro se encuentra igualmente apareado con el nucleótido **+2** de la misma hebra.



Este primer enlace corresponde, por tanto, al extremo 5'-P del RNA naciente. La reacción se repite sobre el extremo 3' del dinucleótido resultante, y así sucesivamente hasta formar un oligorribonucleótido de unas 10 bases, que permanece temporalmente unido a la hebra molde de DNA como híbrido RNA-DNA. En este momento, la RNAPol-II se separa de los TFs unidos a los promotores (*liberación del promotor*) y comienza a desplazarse por el DNA (en sentido 3'→5' de la hebra molde). Este momento se considera el final de la etapa de iniciación y el comienzo de la de elongación.

19.4.1.4 Liberación del promotor: transición de iniciación a elongación

En **procariontes**, esta transición se desencadena por la disociación de la subunidad σ de la holoenzima RNAPol, quedando aquella disponible para su interacción con una nueva RNAPol núcleo e iniciar la transcripción en un nuevo promotor. El resto de la polimerasa (RNAPol "núcleo", desprovista de la subunidad σ) continúa actuando en la elongación.



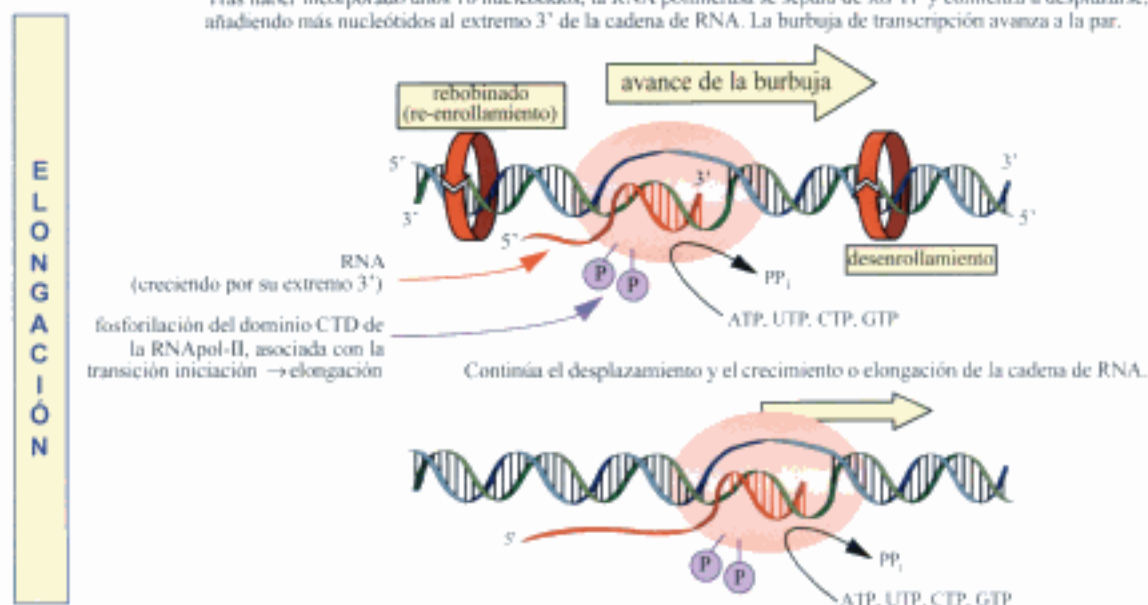
En **eucariotas**, la transición entre las fases de iniciación y elongación viene marcada por la fosforilación del dominio CTD de la RNAPol-II (en la fase de iniciación participó la forma no fosforilada). Esta fosforilación la lleva a cabo el factor de transcripción TFIIF.

19.4.2 Elongación o alargamiento del RNA

La RNA polimerasa continúa alargando la cadena de RNA mientras avanza por el DNA, desplazando junto con ella la burbuja de transcripción. El complejo de elongación ya está sólo formado por el DNA abierto, la RNAPol y el RNA naciente. Durante este avance continuo ocurre lo siguiente:

- Por delante de la polimerasa se separan las dos hebras del DNA por ruptura de los puentes de hidrógeno (avance de la burbuja). La tensión creada por el efecto de superenrollamiento debe ser compensada por topoisomerasas (págs. 66 y 151).
- La polimerasa alarga el RNA añadiendo nucleótidos al extremo 3' del RNA naciente. El último tramo permanece dentro de la burbuja, apareado transitoriamente como híbrido RNA-DNA.
- Por detrás de la polimerasa, la cadena de RNA nuevo se va separando de la hebra molde de DNA, saliendo de la burbuja y quedando como RNA de hebra sencilla. En el DNA se vuelven a aparear las dos hebras y se produce el re-enrollamiento o rebobinado de la doble hélice, de nuevo asistido por topoisomerasas.

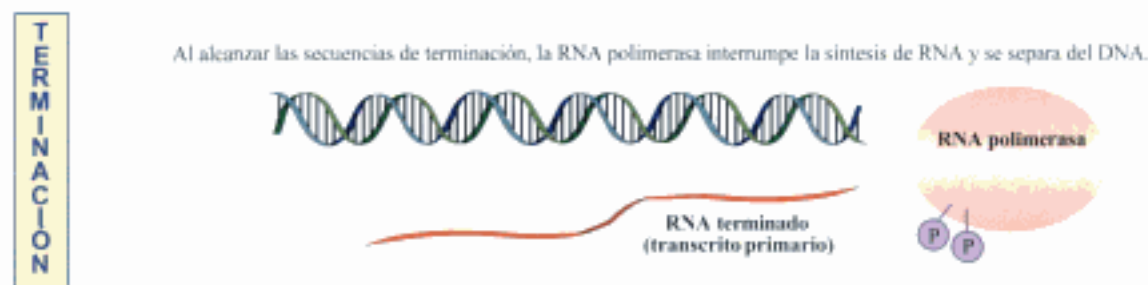
Tras haber incorporado unos 10 nucleótidos, la RNA polimerasa se separa de los TF y comienza a desplazarse, añadiendo más nucleótidos al extremo 3' de la cadena de RNA. La burbuja de transcripción avanza a la par.



Para que la elongación sea eficiente se requiere, además, la ayuda de los llamados **factores de transcripción generales de la elongación** (PTE). Se desconoce su mecanismo exacto de actuación. Son proteínas que se agrupan en varios tipos: PTEFb y SII evitan la terminación prematura de la elongación, mientras que FIIF, SIII (o elonguina) y ELL aumentan la velocidad de la elongación al suprimir las pausas transitorias.

19.4.3 Terminación

El complejo de transcripción responde a señales específicas para finalizar el proceso, separándose el RNA formado.



19.4.3.1 Terminación en procariotas

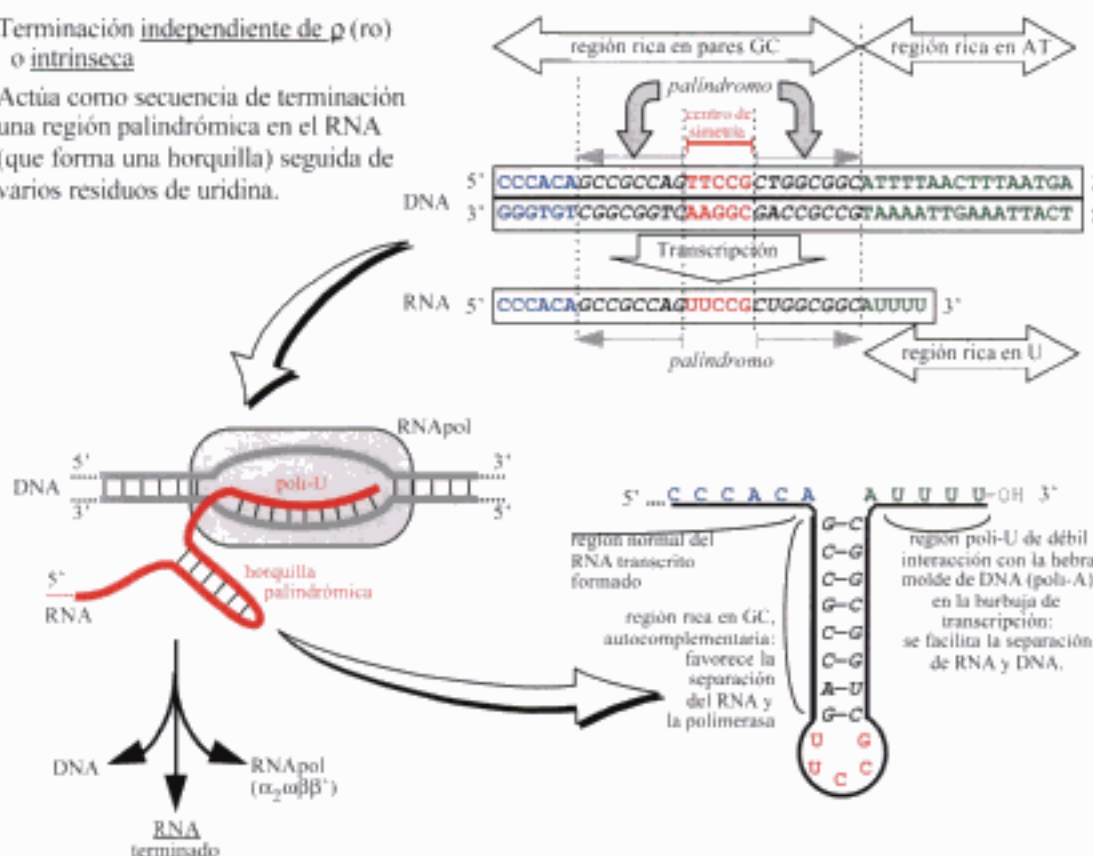
Es importante describir los dos mecanismos básicos de terminación de la transcripción, bien diferenciados en procariotas, y que han servido de soporte para comprender los mecanismos más complejos en eucariotas.

El primero, denominado **terminación independiente de ρ** , se caracteriza por la presencia en el RNA en formación, en una posición cercana a su extremo 3', de un "punto de terminación" definido por una secuencia palindrómica, intrínseca al RNA. Obviamente, esta secuencia se debe a una región palindrómica en el DNA de doble hebra, situado en ese momento en la burbuja de transcripción. Como consecuencia, el RNA forma una horquilla, lo que por alguna razón, no claramente descrita, hace que éste se separe del DNA molde y se interrumpa la síntesis del RNA.

El mecanismo denominado **terminación dependiente de ρ** se basa en la presencia en el RNA de otra secuencia específica, rica en C y pobre en G, de 50-90 nt de longitud. Cuando esta secuencia está situada en el RNA ya formado y fuera de la burbuja, interacciona con una proteína llamada ρ (ro), hexamérica (275 kDa en total), con actividades DNA/RNA-helicasa y ATPasa dependiente de RNA. Al parecer, una vez unida ρ empieza a desplazarse sobre el RNA en sentido 3', hasta alcanzar el complejo de elongación. En este contacto, ρ provoca la desnaturalización del híbrido, liberando el RNA e interrumpiendo la transcripción.

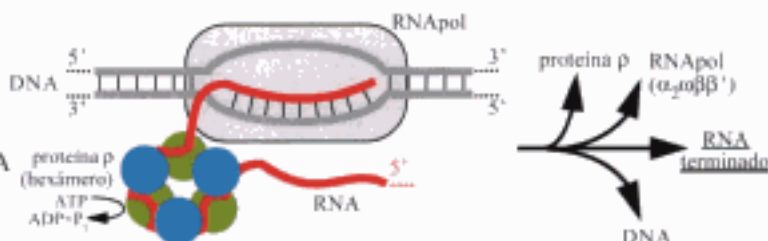
Terminación independiente de ρ (ro) o intrínseca

Actúa como secuencia de terminación una región palindrómica en el RNA (que forma una horquilla) seguida de varios residuos de uridina.



Terminación dependiente de ρ (ro)

La proteína ρ se une a secuencias específicas del RNA, favoreciendo su separación del DNA y, en consecuencia, el fin de la transcripción.



19.4.3.2 Terminación en eucariotas

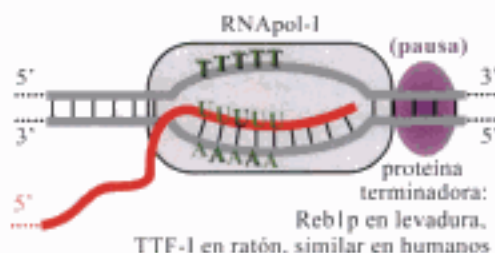
Al alcanzar determinadas secuencias del DNA eucariótico, aún poco caracterizadas, se produce la terminación, como consecuencia de la combinación de dos factores o mecanismos: la **detención o pausa** de la RNA polimerasa y la **desestabilización del híbrido** RNA/DNA. El resultado final es la separación de los componentes del complejo de transcripción y, en el caso de la RNAPol-II, su desfosforilación para que pueda ensamblarse a un nuevo complejo de iniciación y comenzar otro ciclo de transcripción.

La influencia de los dos mecanismos mencionados es diferente para cada una de las tres polimerasas eucarióticas. A diferencia de procariotas, parece ser un hecho común la ausencia de regiones palindrómicas del RNA que participen en la terminación de la transcripción. Las características para cada RNAPol son:

Terminación de RNAPol-I en eucariotas:

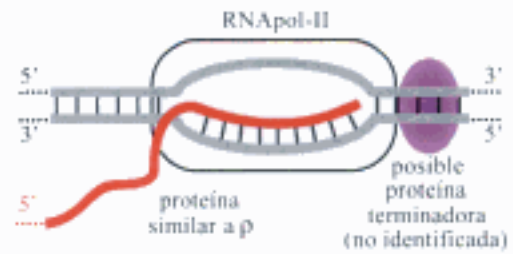
mecanismo similar a la terminación intrínseca de procariotas.

- Una "proteína de pausa" o "proteína terminadora" se une a secuencias específicas del DNA y detiene el avance de la RNAPol.
- Una vez detenida la polimerasa, la presencia de una región rica en T en la hebra no-molde facilita la separación del RNA.
- No hay región palindrómica.



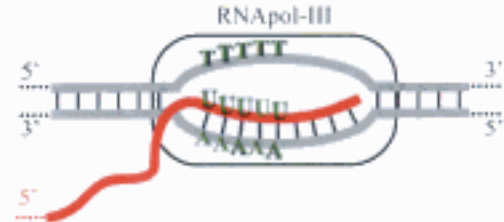
Terminación de RNAPol-II en eucariotas:
mecanismo similar al dependiente de ρ en procariotas.

- En múltiples sitios dentro de una amplia "región de terminación" la RNAPol se detiene, lo que puede deberse a secuencias específicas en el DNA o a "proteínas de pausa" que se unen al DNA.
- Una proteína, similar a ρ (de hecho, la propia ρ bacteriana produce el mismo efecto *in vitro*), se une al RNA induciendo su separación.
- No hay región palindrómica.



Terminación de RNAPol-III en eucariotas:
mecanismo similar a la terminación intrínseca de procariotas.

- La RNAPol se detiene en secuencias específicas del DNA (no se han encontrado proteínas terminadoras).
- Una vez detenida la polimerasa, la presencia de una región rica en T en la hebra no-molde facilita la separación del RNA.
- No hay región palindrómica

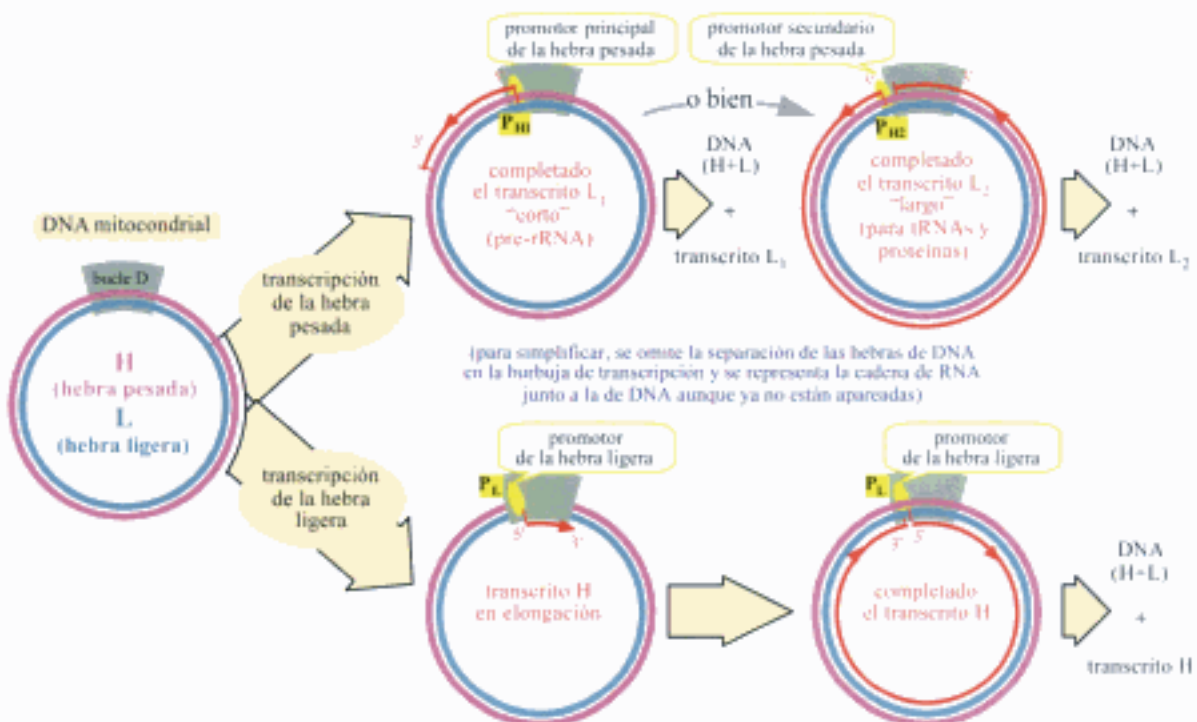


Finalmente, e independientemente de estos mecanismos, es importante adelantar que el extremo 3' del RNA funcional no corresponde al punto donde termina la transcripción, debido a que durante la maduración postranscripcional se corta la molécula de RNA recién transcrita (pág. 279). Por esta razón es menos importante precisar la terminación en eucariotas que en procariotas, y también se explica la dificultad del estudio de las secuencias y mecanismos de terminación eucarióticos (pues es difícil aislar los transcritos primarios, de corta vida).

19.5 TRANSCRIPCIÓN DEL GENOMA MITOCONDRIAL

Como ya se ha indicado, el cromosoma circular mitocondrial humano, y de mamíferos en general, es transcrito por una sola RNAPol que está formada por un solo polipéptido y que, para la iniciación, requiere la presencia de un factor de transcripción mtTFA. El mtDNA se transcribe a partir de 3 lugares de iniciación, dos para la hebra pesada (H1 y H2) y uno para la ligera (L), produciéndose así tres transcritos primarios diferentes (dos de ellos complementarios y correspondientes a toda la longitud del cromosoma, 16,5 kb). Los 3 promotores se encuentran situados en la región no codificante de control o **bucle D** (págs. 160 y 378). Aún se desconoce si el inicio en H1 y H2 depende de un mismo promotor o de dos independientes ("principal" y "secundario"), uno de ellos todavía desconocido.

A partir del primer lugar de iniciación de la hebra pesada (H1) se transcribe el DNA que codifica los tRNA^{Phe}, rRNA 12S, tRNA^{Val} y rRNA 16S; la transcripción se detiene tras este último, mediante la participación de un factor de terminación (mTERF). Cuando la transcripción comienza en el segundo punto de inicio H2, localizado inmediatamente tras el tRNA^{Phe}, el transcrito primario (L2) continúa más allá del punto de terminación anterior, completando el resto del



cromosoma. Por tanto, el lugar de iniciación define qué tipo de RNAs, ribosómicos o mensajeros, deben sintetizarse; la selección del punto de inicio está a su vez regulada por las hormonas tiroideas. Gracias a esta doble unidad de transcripción la mitocondria sintetiza mucha mayor cantidad de RNAs ribosómicos que del resto de productos génicos (unas 60 veces más). (Compárese con la situación análoga en el genoma nuclear, conseguida gracias a la abundancia de la RNAPol-I, especializada en transcribir rRNA, pág. 251.)

A partir del promotor de la hebra ligera se sintetiza una sola molécula (que es una hebra pesada, H) que sólo codifica 8 tRNAs y un mRNA.

Los 3 transcritos primarios sufren luego maduración para dar lugar cada uno a varios rRNAs, tRNAs y mRNAs maduros (pág. 284); del transcrito H se origina también el RNA cebador para la replicación de la hebra H.

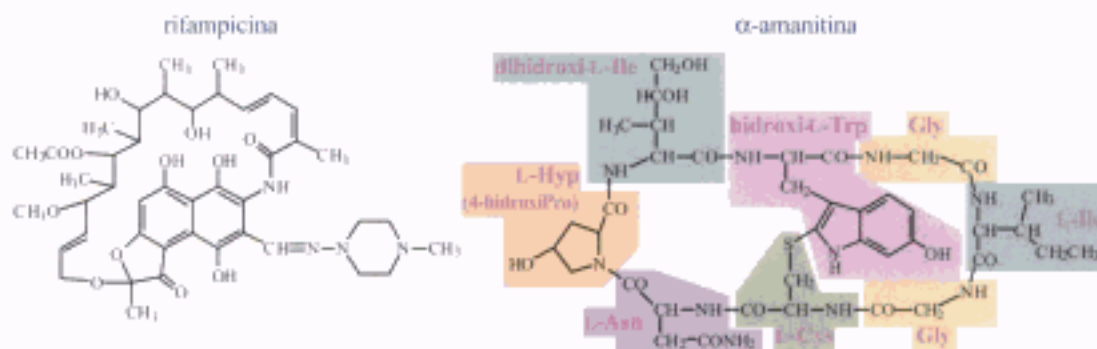
19.6 INHIBIDORES DE LA TRANSCRIPCIÓN

Existe una serie de compuestos que inhiben la transcripción al actuar sobre la RNA polimerasa a través de distintos mecanismos. Son de interés desde el punto de vista experimental y por su aplicación como antibióticos, por mostrar algunos distinta actividad sobre las células procarióticas y eucarióticas.

19.6.1 Antibióticos de tipo I

Este grupo está formado por los compuestos que se unen a la RNA polimerasa, inactivándola de modo directo. Pertenecen a este grupo:

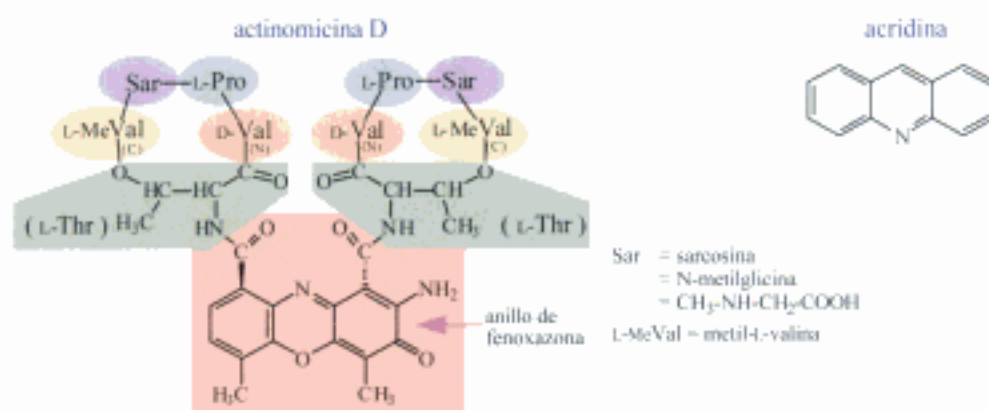
- **Rifamicinas** (producidas por *Streptomyces mediterraneus*) y sus derivados semisintéticos (como la rifampicina), que se unen específicamente a la subunidad β de la RNAPol bacteriana, impidiendo el inicio de la transcripción. La permanencia de la enzima inactivada sobre el promotor impide la unión de otras moléculas de RNAPol no modificadas. Sin embargo, no tienen efecto sobre las polimerasas nucleares eucarióticas, por lo que se utilizan como agentes bactericidas frente a diversos tipos de infecciones. La RNAPol mitocondrial también es inhibida, aunque con menor potencia que la procariótica.
- **Estreptolidigina** (producida por *Streptomyces lydicus*), que también se fija a la subunidad β , pero inhibe la elongación.



La α -amanitina es un inhibidor específico de la etapa de elongación de la síntesis de RNA en eucariotas y no en procariotas. Su efecto inhibitorio se ha empleado como un aspecto diferenciador entre las tres RNAPol eucarióticas (pág. 251). Este agente tóxico es un mecanismo de defensa de la seta venenosa *Amanita phalloides*, para evitar su ingestión. La intoxicación no provoca síntomas inmediatos, sino el deterioro de la función hepática en los días siguientes, debido a la falta de producción de nuevos mRNAs para la síntesis continua de proteínas.

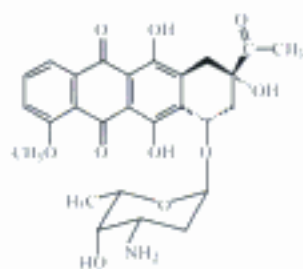
19.6.2 Antibióticos de tipo II

En este caso, la transcripción se inactiva indirectamente, pues el antibiótico se une al DNA impidiendo la unión de la polimerasa o bien su avance. Por ello, se inhiben por igual todas las RNAPol, así como las DNAPol (es decir, la replicación). Éste es el caso de la **actinomicina D**, sintetizada por varias especies de *Streptomyces*, formada por dos pentapéptidos cíclicos unidos a un sistema anular plano de fenoxazona. Éste se intercala en la doble hélice del DNA, entre dos pares de bases G=C sucesivos, deformando el DNA. Al mismo tiempo, las dos estructuras pentapéptidicas cíclicas abrazan a la doble hélice, situándose en el surco menor. Esta alteración impide el movimiento de la DNAPol a lo largo del molde. La actinomicina D es muy tóxica, pues impide la formación de los complejos abiertos, bloqueando el paso de cualquier tipo de polimerasa. Al inhibir la elongación en células intactas y en extractos celulares, es muy útil para identificar procesos celulares que dependen de la síntesis de RNA.

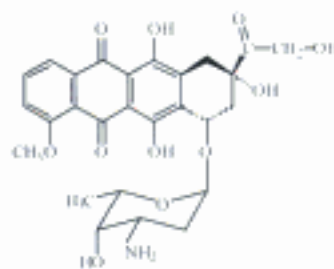


Otros compuestos intercalantes son la **acridina** y sus derivados, que inhiben la síntesis de RNA de una forma similar, y la **daunomicina** y sus derivados, que se emplean como quimioterapéuticos en diversos tipos de cáncer, ya que actúan preferentemente sobre células con crecimiento rápido. Por último, el **bromuro de etidio**, de uso común en el laboratorio para la tinción de DNA en electroforesis (pág. 138), es también un agente intercalante (aunque no empleado como antibiótico), lo que justifica su toxicidad, que debe tenerse en cuenta durante su manipulación.

daunomicina o daunorrubicina



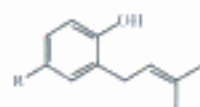
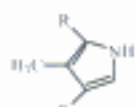
doxorubicina o adriamicina



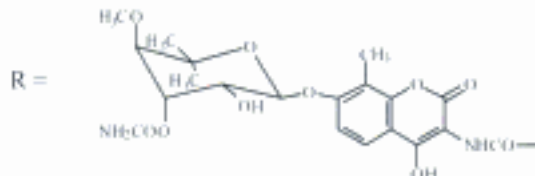
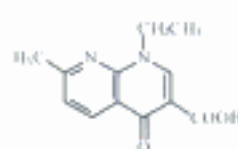
19.6.3 Antibióticos de tipo III

Los antibióticos de tipo III son también inhibidores indirectos, que actúan en este caso impidiendo la acción de la **DNA girasa** procariótica, la enzima que introduce superenrollamientos negativos en el DNA (pág. 152). Se comprueba así que éstos son necesarios para la separación de las hebras de DNA, lo que permite la replicación y la transcripción. Los principales antibióticos en este grupo son la **cumermicina**, la **novobiocina** y el **ácido nalidíxico**.

novobiocina

cumermicina A₁

ácido nalidíxico



Tema 20

Control de la expresión génica: pretranscripcional y transcripcional

20.1	Introducción general a la regulación de la expresión génica.....	261
20.2	Control pretranscripcional.....	263
20.3	Regulación genética de la transcripción.....	264
20.3.1	Elementos que actúan en <i>cis</i> y <i>trans</i> como reguladores de la transcripción.....	264
20.3.2	Promotores de los genes de clase II.....	265
20.3.2.1	Secuencias o elementos basales del promotor.....	265
20.3.2.2	Secuencias o elementos proximales.....	265
20.3.2.3	Secuencias o elementos distales.....	266
20.3.3	Factores de transcripción de clase II.....	266
20.3.3.1	Factores de transcripción generales.....	267
a)	Ensamblaje consecutivo de los factores generales (modelo "escalonado").....	267
b)	Modelo de la holoenzima.....	267
20.3.3.2	Factores de transcripción proximales.....	269
20.3.3.3	Factores de transcripción inducibles.....	269
20.3.3.4	Interacción de los factores de transcripción con la polimerasa.....	270
20.3.3.5	Iniciación de la transcripción a partir de diferentes promotores.....	270
20.3.4	Regulación de la transcripción de genes de clases I y III.....	271
20.4	Regulación epigenética.....	271
20.4.1	Metilación del DNA y control de la expresión.....	272
20.4.2	Patrón de metilación durante el desarrollo.....	272
20.4.3	Metilación y cáncer.....	273

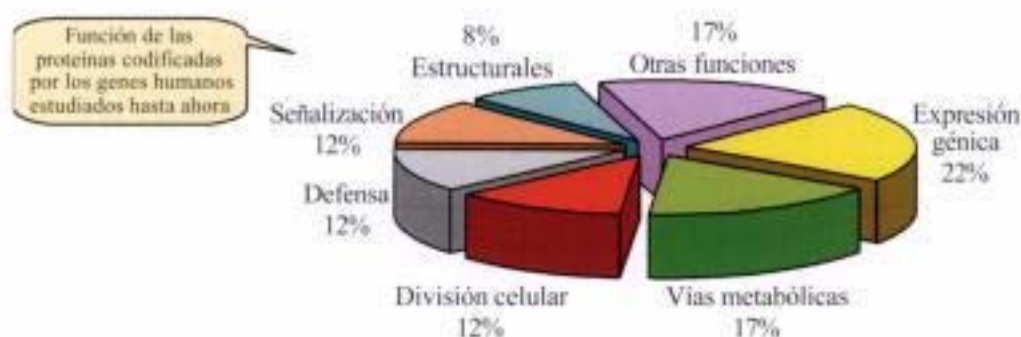
20.1 INTRODUCCIÓN GENERAL A LA REGULACIÓN DE LA EXPRESIÓN GÉNICA

La visión global del control de la expresión génica es especialmente difícil en eucariotas en general, y en organismos pluricelulares en particular. Por ejemplo, aunque todas las células humanas contienen el mismo genoma (salvo mutaciones somáticas), existen claras diferencias frente a las células procarióticas. Es casi un misterio el hecho de que las células de organismos complejos, procedentes del desarrollo de una misma célula (el cigoto) y conteniendo un DNA idéntico, sean capaces de expresar funciones tan diferentes y únicas en cada órgano o tejido particular. Esta enorme flexibilidad en el control de sus genes viene determinada por los siguientes tipos de *expresión génica diferencial*:

- Distintos tipos de células en cada especie y, dentro de éstas, en cada órgano o tejido. En humanos, se calculan unos 200 tipos celulares, agrupados en "familias" según su función específica: células epiteliales, contráctiles, sanguíneas, del sistema inmunitario, sensoriales, neuronales, etc.
- Variedad en el número de genes en cada especie. En humanos, el total de genes parece ser del orden de 75.000 a 100.000, aunque recientemente se dan cifras aún más discordantes, entre 30.000 y 200.000 genes.
- Proporción de genes expresados en cada tipo celular. En humanos, una célula sólo expresa en cada momento un 10-20% del total de genes; es decir (asumiendo 50.000-100.000 genes), sólo se expresan entre 500 y 2.000 genes en cada célula.
- Influencia sobre la expresión de cada gen de los estadios del ciclo de división celular y del desarrollo del individuo, necesidades metabólicas o de otro tipo, respuesta a agentes externos, etc.

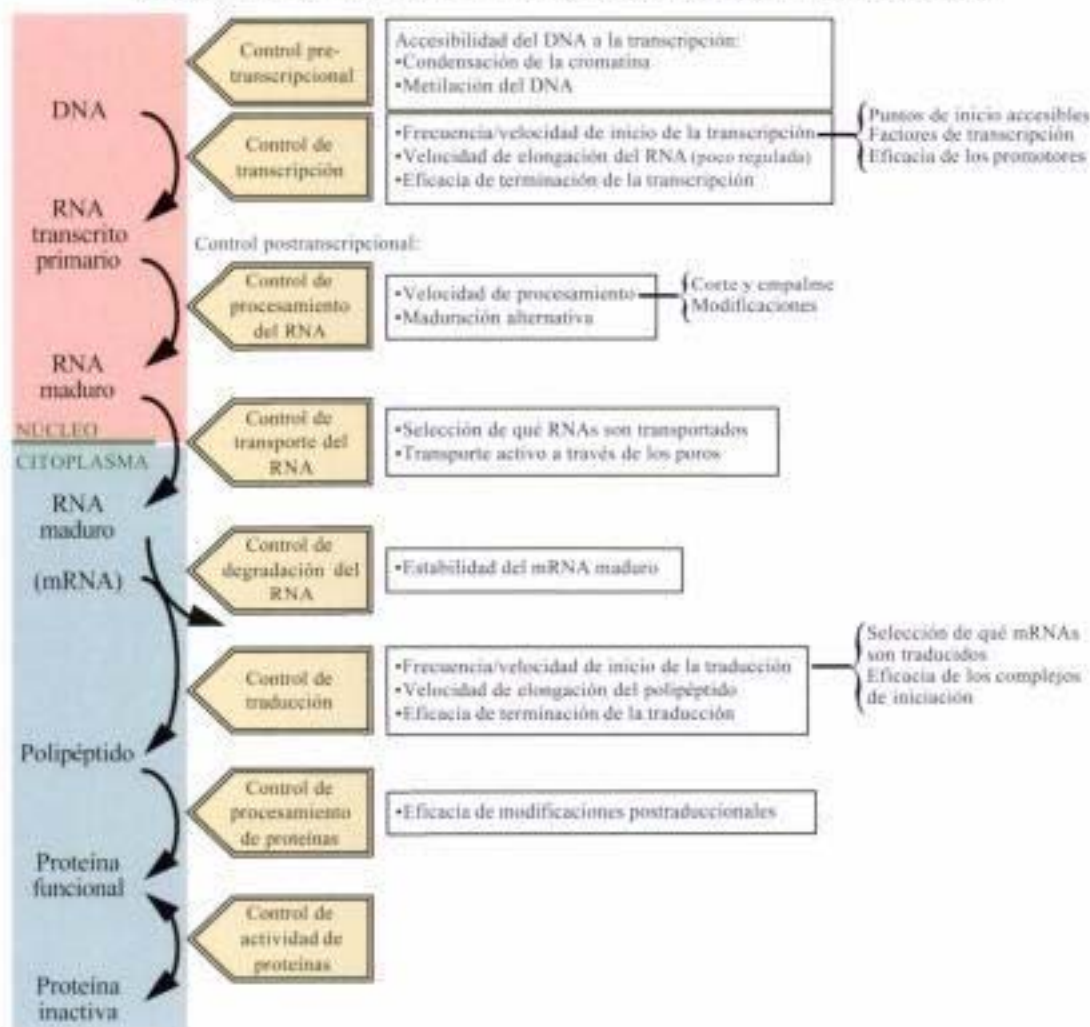
Dentro de estas posibilidades, es de destacar la variación de la expresión génica durante el proceso de diferenciación celular, es decir, la adquisición de propiedades específicas por parte de las diversas células de un organismo, que tiene lugar durante la embriogénesis. Ello hace que existan genes que no responden a cambio fisiológico alguno o que, por el contrario, estén sometidos a fuerte control como consecuencia del desarrollo animal o vegetal, de la organización de células en tejidos y de tejidos en organismos completos.

La manifestación principal de esta flexibilidad en la expresión génica es la aparición en cada tipo celular de proteínas en proporciones diversas, desde muy elevadas en unos casos a muy bajas en otros. Dado que las proteínas condicionan las propiedades celulares (arquitectura celular, actividades enzimáticas, propiedades fisiológicas, interacción con el ambiente, etc.), todo lo anterior quiere decir que es variable la proporción de genes implicados en la formación de proteínas para una determinada "tarea" o función:



La expresión de los genes de un organismo superior puede ser regulada a distintos niveles. Desde el punto de vista patológico, es importante resaltar que la alteración de cualquiera de los pasos implicados en la expresión génica pueden causar enfermedades.

NIVELES DE REGULACIÓN DE LA EXPRESIÓN GÉNICA EN EUCARIOTAS



El estudio en profundidad de todos estos niveles de control debe hacerse en tratados especializados. En este texto sólo trataremos de desglosar de forma paulatina lo esencial, cada uno de ellos en su correspondiente proceso de expresión génica (temas 21, 23 y 24).

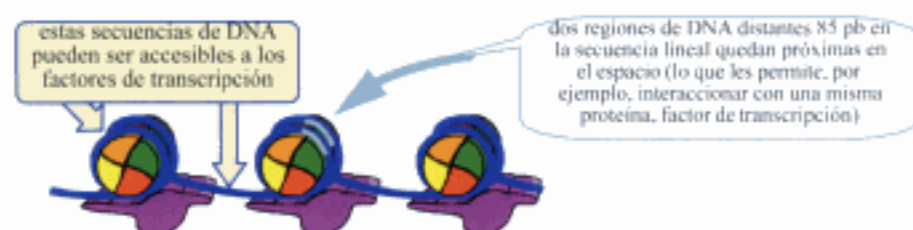
20.2 CONTROL PRETRANSCRIPCIONAL

El inicio de la transcripción requiere que la maquinaria molecular responsable (RNA polimerasa y factores de transcripción, pág. 264) pueda acceder físicamente a las bases de la hebra de DNA que se ha de transcribir. Ello supone la descondensación previa del cromosoma durante la interfase del ciclo de división celular (pág. 98), mediante mecanismos moleculares aún no conocidos exactamente.

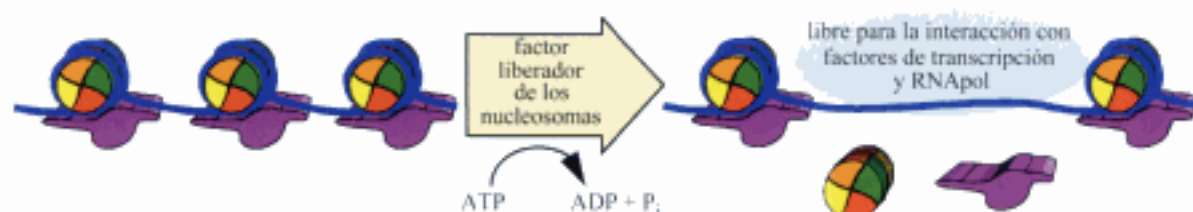
La transcripción no puede tener lugar sobre DNA en estados de condensación superiores a la fibra de 10 nm (rosario de nucleosomas, pág. 85). En ocasiones, el DNA se puede transcribir cuando se encuentra bajo la forma de complejos con las histonas (nucleosomas), pero también puede ser necesario que éstos se desensamblen para permitir el acceso al DNA. En general, la accesibilidad del DNA a la transcripción, o **actividad transcripcional de la cromatina**, está relacionada estrechamente con el grado de condensación (pág. 87) y puede valorarse experimentalmente por la sensibilidad a la acción de DNasas.

Se puede considerar que el acceso al DNA necesario para la transcripción es una combinación de tres mecanismos:

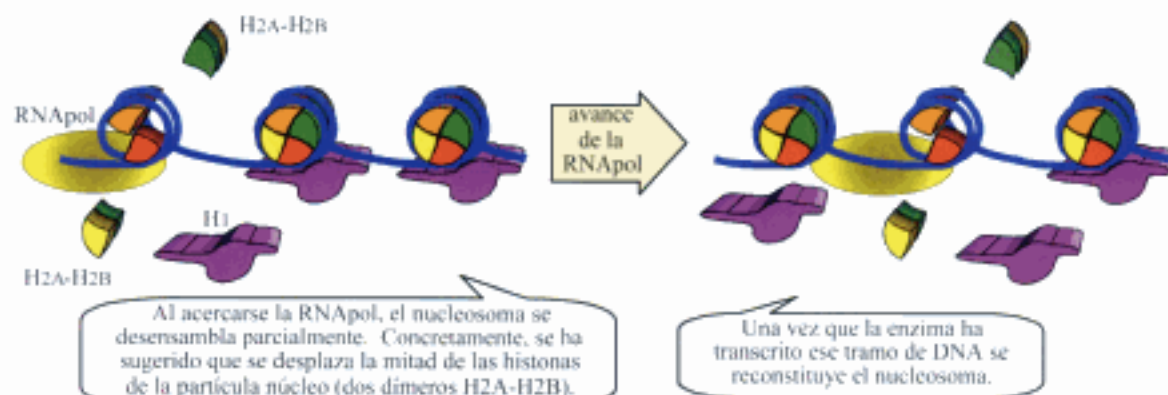
- a) **Posición precisa de los nucleosomas.** El reconocimiento de las secuencias promotoras del DNA por los factores de transcripción (pág. 264) puede tener lugar, al menos en algún caso, gracias a una posición muy específica de los nucleosomas, de forma que tales secuencias queden en la región internucleosómica (DNA espaciador o *linker*, pág. 86) o, si forman parte del nucleosoma, queden orientadas hacia su superficie externa.



- b) **Retirada de los nucleosomas.** Tanto el reconocimiento de las secuencias promotoras por los factores de transcripción como la unión y avance de la RNApol pueden requerir la liberación de las histonas, o al menos el desplazamiento de los nucleosomas, para hacer el DNA accesible. Se han encontrado proteínas capaces de efectuar este "remodelado de la cromatina", gracias a la hidrólisis de ATP.



- c) **Avance compatible con la presencia de nucleosomas.** La transcripción puede tener lugar en regiones de DNA con nucleosomas íntegros; posiblemente, la RNA polimerasa se desplaza a lo largo de la fibra de 10 nm, transcribiendo una hebra gracias a un desensamblado parcial y reversible de los nucleosomas.



propia RNAPol en el reconocimiento de la secuencia del promotor y se encargan de la regulación de la transcripción. Su diversidad es aún mayor que la de promotores, lo que permite la gran flexibilidad existente en el control de la expresión de los genes. Los TFs se denominan también "**factores que actúan en trans**" o "**factores trans**", pues sus genes están en posición alejada y no relacionada con aquella región génica cuya expresión regulan. Su capacidad de unión a la molécula de DNA, casi siempre reconociendo ciertas secuencias de bases, está generalmente mediada porque la presencia en las moléculas de TF de motivos estructurales de unión al DNA, como dedos de zinc, homeodomios, etc. (pág. 60).

La especialización de las RNAPol eucarióticas para transcribir sólo determinados genes (pág. 251) viene determinada por la regulación de estos genes por secuencias promotoras y factores de transcripción diferentes. Antes de describir los pormenores del control de la transcripción de los genes de clase II conviene recopilar la terminología empleada para las moléculas que intervienen. Se distingue así entre promotores para genes de la clase I, de la clase II y de la clase III. Los factores de transcripción respectivos se nombran con las iniciales TF seguidas del tipo de RNAPol con la que pueden actuar (I, II o III) y una letra que los identifica (por ejemplo, TFIIA, TFIIB, TFIID son 3 factores de transcripción para la RNAPol-II).

polimerasa (enzima)	genes que transcribe (DNA)	producto génico (RNA o proteína)	promotores (DNA)	factores de transcripción (proteínas)
RNAPol-I	"de clase I"	rRNAs de 28S, 18S y 5.8S	"de clase I"	TFI
RNAPol-II	"de clase II"	proteínas y snRNAs	"de clase II"	TFII
RNAPol-III	"de clase III"	tRNAs, rRNA-5S y RNAs pequeños	"de clase II"	TFIII

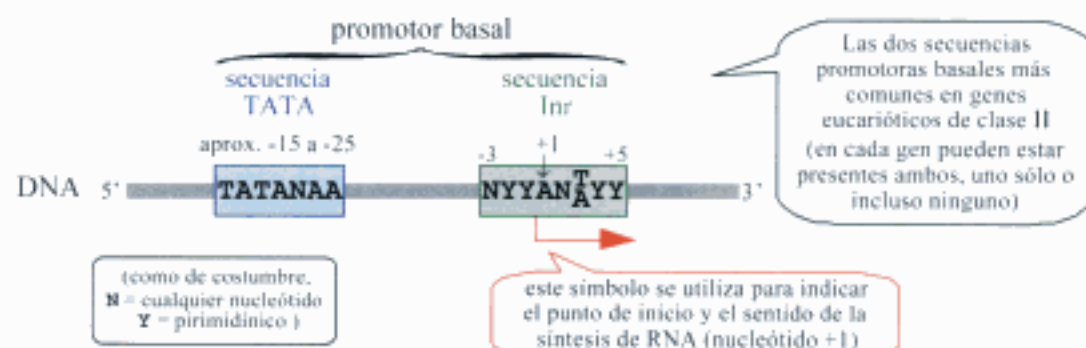
20.3.2 Promotores de los genes de clase II

La regulación de los genes de clase II (transcritos por la RNAPol-II) es del máximo interés por ser responsable de la síntesis de los mRNA y, por tanto, de la síntesis de todas las proteínas. Sus promotores se localizan típicamente en dirección 5' (corriente arriba) del origen de transcripción y muestran una mayor variación de secuencia que los promotores para RNAPol-I y III.

Los promotores en general, pero de forma característica los de clase II, se pueden dividir en 3 tipos, de acuerdo con su acción y su ubicación. Su función se comprenderá mejor al estudiar los factores de transcripción que se unen a ellos.

20.3.2.1 Secuencias o elementos basales del promotor

El **promotor basal** comprende las secuencias que definen el punto de inicio de la transcripción y son imprescindibles para que ésta comience. Es una región situada en posición adyacente al origen de transcripción, comúnmente corriente arriba hasta la posición -30. Dentro de esta región puede distinguirse como más habitual la **secuencia TATA**, **caja TATA** o **caja de Hogness**, situada típicamente alrededor de las posiciones -15 a -25 (no debe confundirse con la caja TATA de procariotas o caja de Pribnow, con secuencia consenso TATAAT y situada en -10). Otra secuencia basal frecuente es la **secuencia iniciadora Inr**, situada sobre el propio origen, entre -3 y +5.

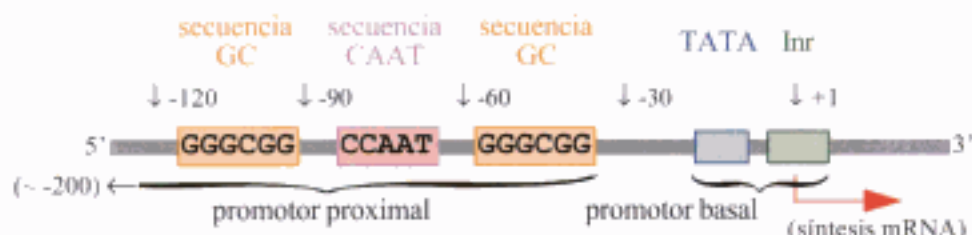


En ocasiones, el promotor de un gen se califica, de acuerdo con las secuencias basales que presenta, como de tipo TATA+Inr+, TATA+Inr-, TATA-Inr+ o TATA-Inr-.

20.3.2.2 Secuencias o elementos proximales

Generalmente, el promotor basal no es suficiente por sí mismo para provocar el inicio de la transcripción, sino que se requiere la intervención adicional de otra región conocida como promotor proximal. Éste está cercano al promotor basal, pero más alejado corriente arriba del origen, comúnmente entre las posiciones -30 y -200. Por tanto, esta región

de DNA suele ser de mayor tamaño que la del promotor basal. Los elementos proximales no especifican la posición de inicio, sino que *determinan la frecuencia con la que se produce el inicio de la transcripción*. Ello es posible gracias a que sobre ellos se unen diversos factores de transcripción que favorecen la interacción de la RNAPol-II con el DNA en el punto de inicio y su actividad enzimática. Aunque en general las secuencias son más variadas que las basales, las dos más características son la **caja o secuencia CAAT**, entre -60 y -80, y la **caja GC**, que aparece en copias múltiples y con cualquier orientación a ambos lados de la caja CAAT.

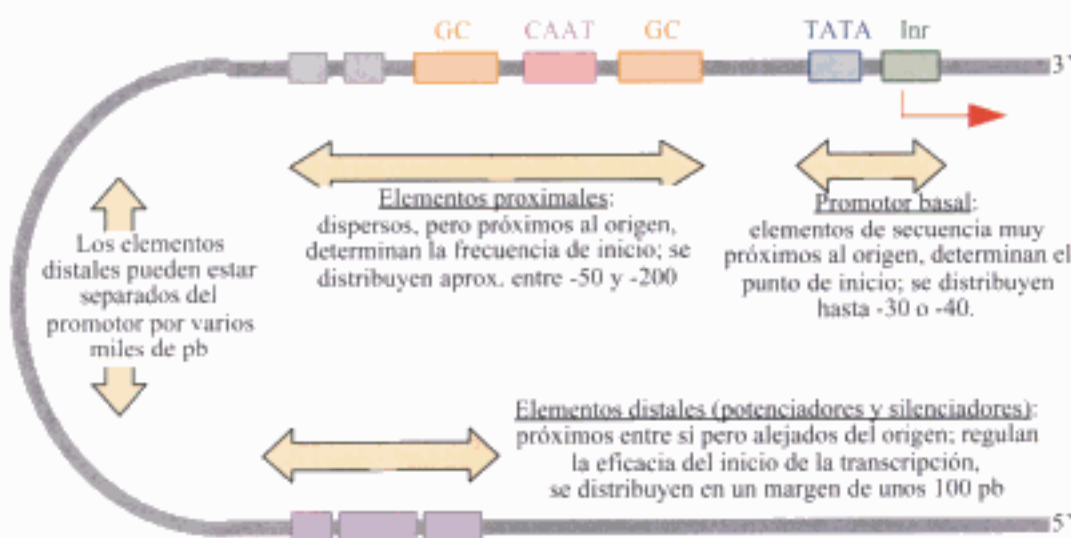


20.3.2.3 Secuencias o elementos distales

La expresión de algunos genes sufre una regulación aún más compleja, que depende de secuencias situadas a gran distancia del punto de inicio, incluso de varios miles de pb, corriente arriba o corriente abajo. Esto ocurre especialmente para los **genes inducibles**, aquéllos que no se expresan continuamente en la célula, sino *cuya expresión sufre una regulación amplia y precisa como respuesta a diversas señales* (por ejemplo, las hormonas esteroideas y tiroideas). Aunque no se suelen contemplar como tales, estas secuencias promotoras deben también considerarse parte del gen, atendiendo al concepto funcional de éste (pág. 244). En general se las llama **secuencias promotoras distales**, por encontrarse alejadas del origen, o también, de acuerdo con su función, **elementos específicos de regulación, módulos de control o elementos de respuesta**.

Estas secuencias promotoras son muy variadas y específicas para cada gen. Otra particularidad es que actúan tanto activando la transcripción como reduciéndola. De acuerdo con ello, se clasifican en dos tipos:

- **Potenciadores**, activadores, intensificadores, amplificadores o estimuladores (*enhancers*). Pueden aumentar mucho la velocidad de inicio de la transcripción (y, por consiguiente, la del proceso completo) conseguida a partir de promotores basales y proximales situados en la misma molécula de DNA.
- **Silenciadores** o inhibidores (*silencers*). Tienen el efecto opuesto, generalmente porque los factores de transcripción que se unen a ellos compiten con la acción de los específicos de los promotores potenciadores y proximales.



20.3.3 Factores de transcripción de clase II

La RNAPol-II interactúa con gran variedad de factores de transcripción (TFII) dependiendo del gen, habitualmente de forma simultánea con un número considerable de ellos. Estos factores se pueden clasificar en generales, proximales e inducibles, de acuerdo con su unión a uno de los 3 tipos de secuencias o elementos promotores recién estudiados. Se estudiarán a continuación con cierto detalle los que actúan sobre el promotor TATA, mejor conocidos.

20.3.3.1 Factores de transcripción generales

Reciben este nombre los TFs que reconocen los elementos basales de un promotor. Son proteínas muy conservadas evolutivamente, que promueven la formación alrededor del punto de inicio de un complejo plurimolecular, que incluye el promotor basal, los propios TF generales y la RNAPol-II. Se puede decir que al reconocer el promotor basal actúan de mediadores para que se fije la polimerasa, definiendo así el punto de inicio de la transcripción y activando a la enzima para que comience a sintetizar el RNA.

Se conocen las funciones de algunos de los factores de transcripción generales asociados a la RNAPol-II:

- **TFIID**, el más significativo de todos ellos, es el único con especificidad por una secuencia de DNA, la de la caja TATA. A pesar de ello, parece actuar también sobre promotores que no tienen caja TATA. Este factor determina que la RNAPol-II actúe a una cierta distancia del promotor, definiendo el punto de inicio de la transcripción (nucleótido +1). Tiene un gran tamaño (masa molecular cercana a 800 kDa), al estar formado por dos tipos de componentes:
 - Una molécula de **TBP**, o "proteína de unión a TATA" (*TATA Binding Protein*), que confiere a TFIID la capacidad de reconocimiento de la secuencia promotora. Su unión a ésta es el primer paso en la formación del complejo de inicio, con la particularidad de unirse al surco menor del DNA (a diferencia de la mayoría de proteínas, que se unen al DNA en el surco mayor) y provoca la flexión de la molécula de DNA.
 - Varias moléculas de **TAFs** o "factores asociados a TBP" (*TBP-associated factors*), diferentes entre sí y que pueden variar para la unión a promotores diferentes. Son subunidades de 30 a 250 kDa que forman parte del TFIID en número de 9 a 12, típicamente.
- **TFIIH** es particularmente importante por su doble actividad enzimática, como helicasa y como proteína quinasa, residentes en distintas subunidades. La helicasa está implicada en la separación de las hebras del DNA en el sitio de inicio, necesaria para el acceso de la polimerasa a las bases, mientras que como quinasa fosforila el dominio C-terminal de la subunidad mayor de la RNAPol-II (CTD, pág. 252), provocando en ésta un cambio conformacional que induce el comienzo de su desplazamiento a lo largo del DNA, la separación de algunos TFs y, en definitiva, la transición entre el inicio y la elongación de la transcripción. Este factor de transcripción es también singular por intervenir en la reparación del DNA por escisión de nucleótidos (pág. 359).

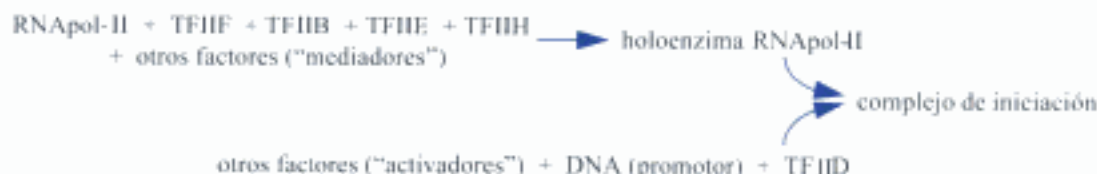
Para la formación del complejo de inicio se han propuesto dos mecanismos, cuyo resultado final es el mismo:

a) *Ensamblaje consecutivo de los factores generales (modelo "escalonado")*

Según esta hipótesis, los TFs y la RNAPol-II se van asociando a la región promotora en un cierto orden uno tras otro (pág. 268); se considera que cada factor va estableciendo un complejo con la estructura adecuada para que se una el siguiente. La formación del complejo de iniciación comienza con la unión de TFIID al DNA en la secuencia TATA y, a continuación, se van asociando los factores TFIIA y TFIIB, que permiten la unión de la polimerasa junto con el factor TFIIE, y finalmente se incorporan TFIIH y TFIIF.

b) Modelo de la holoenzima

En esta hipótesis alternativa, también apoyada por estudios *in vitro*, la RNAPol-II se asocia previamente a varios factores de transcripción, formando una enzima activa u **holoenzima**, que luego se une al DNA en la secuencia promotora definida por la unión del TFIID.

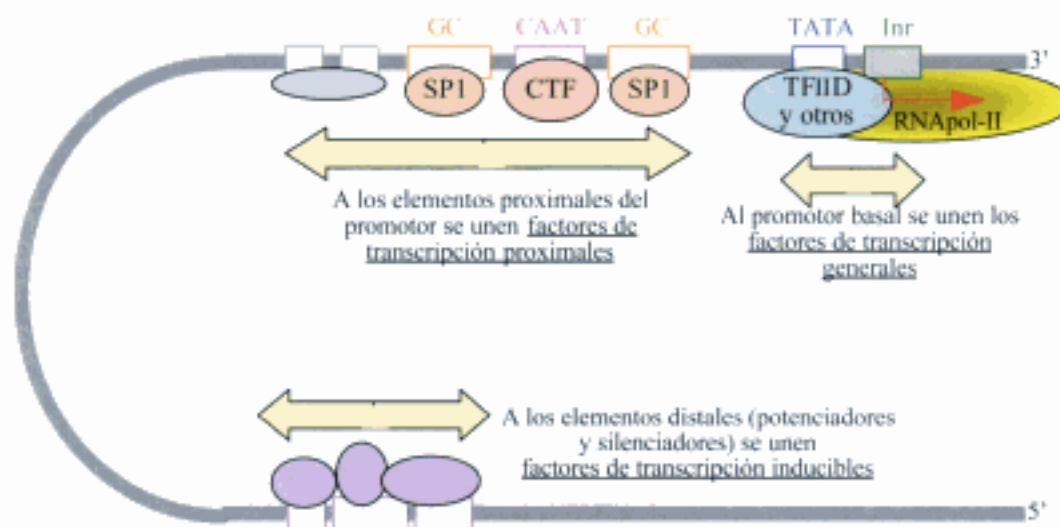


El resultado es el mismo complejo de iniciación que en el modelo escalonado, que permite el comienzo de la transcripción, y la polimerasa luego lo abandonaría y avanzaría igual que se ha indicado en aquel modelo.

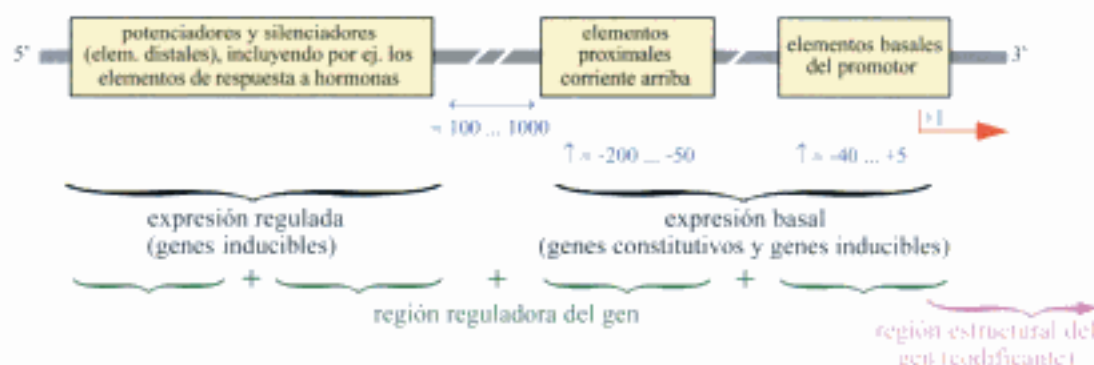
20.3.3.2 Factores de transcripción proximales

También llamados factores de unión corriente arriba, son proteínas que reconocen específicamente los elementos proximales del promotor, contribuyendo a aumentar la eficiencia de formación del complejo de inicio o favoreciendo su actividad sobre la polimerasa. Cada promotor en particular requiere un conjunto preciso de estos factores para su expresión plena.

Como ejemplos, el factor de transcripción **CTF** (o **NF1**) interacciona con la caja CAAT, siendo aparentemente el máximo responsable de la eficiencia del promotor, mientras que el factor de transcripción **SP1** se une a las cajas GC.



Los promotores que sólo contienen elementos reconocidos por TFs generales y proximales controlan los llamados **genes constitutivos**, aquéllos que se expresan de una forma constante, a la misma velocidad en todo momento de la vida celular. Se asume generalmente que este concepto es sinónimo de **genes domésticos o caseros** (*housekeeping genes*), aquéllos que se expresan en todas las células del organismo; el nombre se debe a que codifican los productos que aseguran las funciones indispensables para la vida de la célula. Ambos conceptos incluyen tanto genes que se transcriben con eficiencia, porque sus productos génicos se precisan en forma abundante, como genes que se transcriben a baja velocidad porque sus productos se necesitan en pequeña cantidad, pero siempre de forma continua en ambos casos.



20.3.3.3 Factores de transcripción inducibles

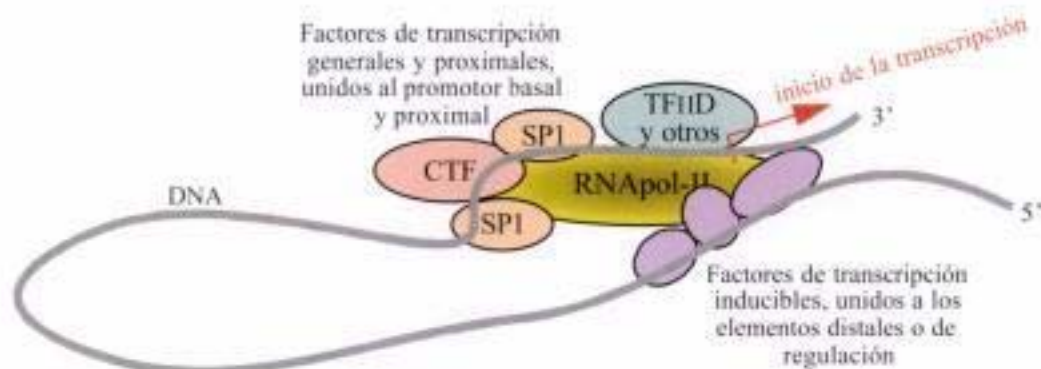
Este tercer grupo de factores de transcripción lo constituyen aquéllos que reconocen los elementos distales del promotor. Allí actúan de forma similar a los factores proximales, es decir, facilitando la formación y actividad del complejo de inicio de la transcripción o, por el contrario, dificultándolas, regulan, por tanto, la transcripción tanto de forma positiva como negativa (de acuerdo con sus promotores, sean potenciadores o silenciadores, pág. 266).

A diferencia de los generales y proximales, estos factores de transcripción no están presentes en la célula de forma continua, sino que se sintetizan o se activan en momentos específicos o en tejidos particulares. Como consecuencia, la expresión de los genes cuya transcripción controlan está igualmente regulada en tiempo y espacio, es decir, se expresan activamente en ocasiones y no lo hacen en absoluto en otras; con frecuencia, en muchos tipos celulares un gen concreto no se expresa nunca. Se los llama por ello **genes inducibles** (por contraposición a los genes constitutivos), **regulables** o **de expresión diferencial**.

Como ejemplo característico se pueden citar los receptores nucleares para hormonas esteroides o tiroideas. La unión de la hormona cambia la conformación del receptor, que actúa como factor de transcripción al aumentar su afinidad por ciertos promotores distales, denominados **elementos de respuesta a hormonas** (HRE). Como consecuencia, se activa o se bloquea, según los casos, la transcripción de genes específicos; los productos proteicos de estos genes representan para la célula el mecanismo de respuesta a la hormona.

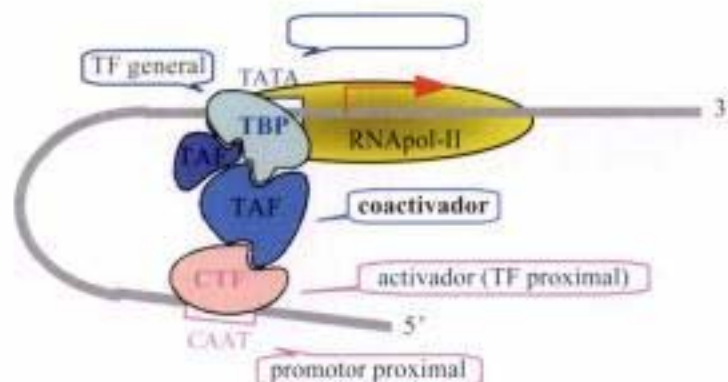
20.3.3.4 Interacción de los factores de transcripción con la polimerasa

Cabe preguntarse cómo los factores que se unen a promotores lejanos al origen de transcripción pueden participar en el complejo de inicio (mecanismos de regulación a distancia). Para explicarlo se debe considerar la flexibilidad de la cadena de DNA, que puede permitir la aproximación física de los tres tipos de secuencias promotoras a pesar de su lejanía en la secuencia lineal de DNA.



De esta forma los factores de transcripción unidos a las regiones basal, proximales y distales pueden contactar entre sí y con la RNAPol-II, por lo que participan al unisono en el control del inicio de la transcripción.

Algunos TFs no interaccionan directamente con la polimerasa, sino que hacen de mediadores entre otros TFs. Se han identificado proteínas o complejos de proteínas que cooperan con los factores activadores, conectándolos con la maquinaria basal para regular la transcripción. Por esta razón, se conocen como **coactivadores** o **cofactores generales**; entre ellos se incluyen los TAFs (pág. 267), que parecen desempeñar un papel de mediadores en la interacción entre los factores generales por un lado y los factores proximales y distales por otro.



20.3.3.5 Iniciación de la transcripción a partir de diferentes promotores

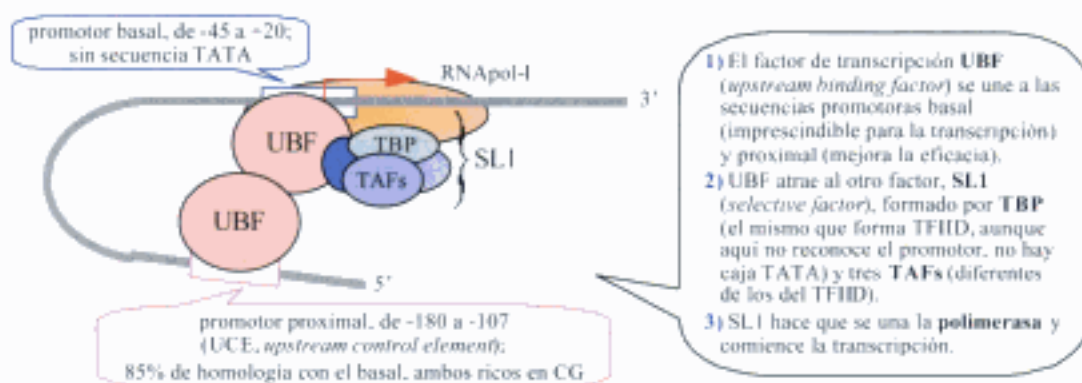
La iniciación de la transcripción en promotores no basados exclusivamente en la caja TATA (pág. 265) tiene lugar con características, aún poco conocidas, diferentes a las estudiadas.

- En los promotores basados en la secuencia Inr pero sin caja TATA (**TATA-Inr+**), la formación del complejo de iniciación tiene lugar cuando se unen al elemento Inr factores de transcripción específicos (**IBP**, proteínas de unión a Inr). Entre éstos se encuentran TAF150, TFII-I, YY1 y la propia RNAPol-II. Por lo demás, el proceso y los requerimientos de unión de factores de transcripción para el inicio son similares a los promotores TATA+ Inr- ya estudiados.
- En los promotores que poseen ambas secuencias (**TATA+Inr+**), entran en juego tanto la TBP como las IBP, dando lugar a la formación de complejos muy estables.
- Existen incluso promotores que no tienen ninguna de las dos secuencias (**TATA-Inr-**); en ellos no se sabe cómo tiene lugar el inicio de la transcripción.

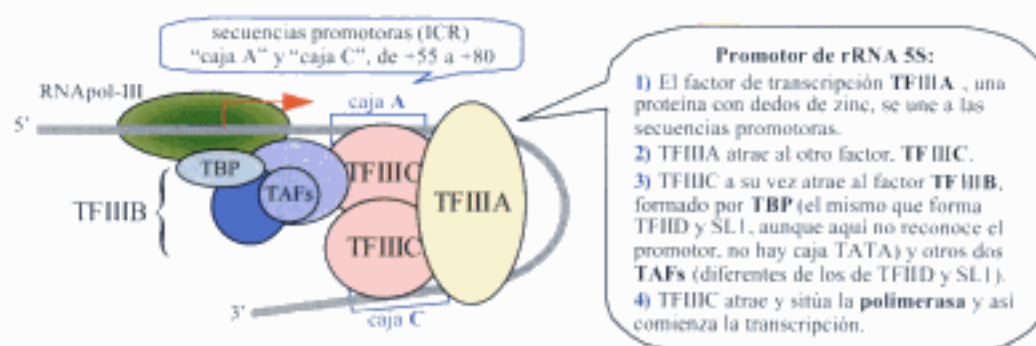
20.3.4 Regulación de la transcripción de genes de clases I y III

Como ya se ha indicado, las secuencias promotoras y el conjunto de factores de transcripción empleados por cada una de las RNA polimerasas eucarióticas son diferentes (pág. 265). La diversidad de promotores y TFs para el inicio por RNAPol-I y III es muy inferior a la correspondiente a la RNAPol-II. Esto puede reflejar simplemente el menor número de genes que son transcritos por aquellas dos enzimas: la RNAPol-I sintetiza un solo transcrito primario (el precursor grande de RNA ribosómico), mientras que la RNAPol-III transcribe unos pocos genes (el pre-rRNA 5S, los cerca de 40 tRNAs y algunos RNA pequeños, pág. 251). Ello contrasta con las decenas de miles de genes proteicos de cuya transcripción es responsable la RNAPol-II, por lo que necesita una *maquinaria* más compleja para su regulación individual, como se acaba de estudiar.

Los promotores de genes de la clase I (para RNAPol-I) se localizan, al igual que los de la clase II estudiados hasta ahora, en dirección 5' (corriente arriba) del punto de inicio.



Entre los genes transcritos por la RNAPol-III se encuentran el de rRNA 5S, los de tRNAs y el de uno de los RNAs nucleares pequeños que intervienen en la eliminación de intrones (snRNA U6; pág. 281); todos ellos tienen promotores y TFs diferentes. Los promotores de rRNA 5S y tRNA aparecen en dirección 3' del inicio (corriente abajo), situándose dentro de la región estructural del gen; por ello se llaman **regiones internas de control (ICR)**. La posición e identidad de las secuencias promotoras y los correspondientes TFs son diferentes para los genes de rRNA 5S, tRNAs y snRNAs; se muestra sólo como ejemplo el primero.



Cabe resaltar que en la iniciación por las 3 RNA polimerasas interviene un factor de transcripción de varias subunidades (**SL1**, **TFIID** y **TFIIIB**), una de las cuales es la proteína de unión a TATA, **TBP**. Las otras subunidades son **TAFs** específicos para cada tipo de promotor. Mientras que en los genes de tipo II **TBP** es el factor inicial que reconoce la caja TATA (pág. 267), esta secuencia no existe en los otros casos (genes de tipo I y III). En ellos, **TBP** no se une directamente al DNA reconociendo la secuencia promotora, sino que requiere otros factores de transcripción para asociarse. Parece que **TBP** desempeña una función común a los 3 tipos de promotores: la interacción con la polimerasa respectiva para situarla en el punto de inicio.

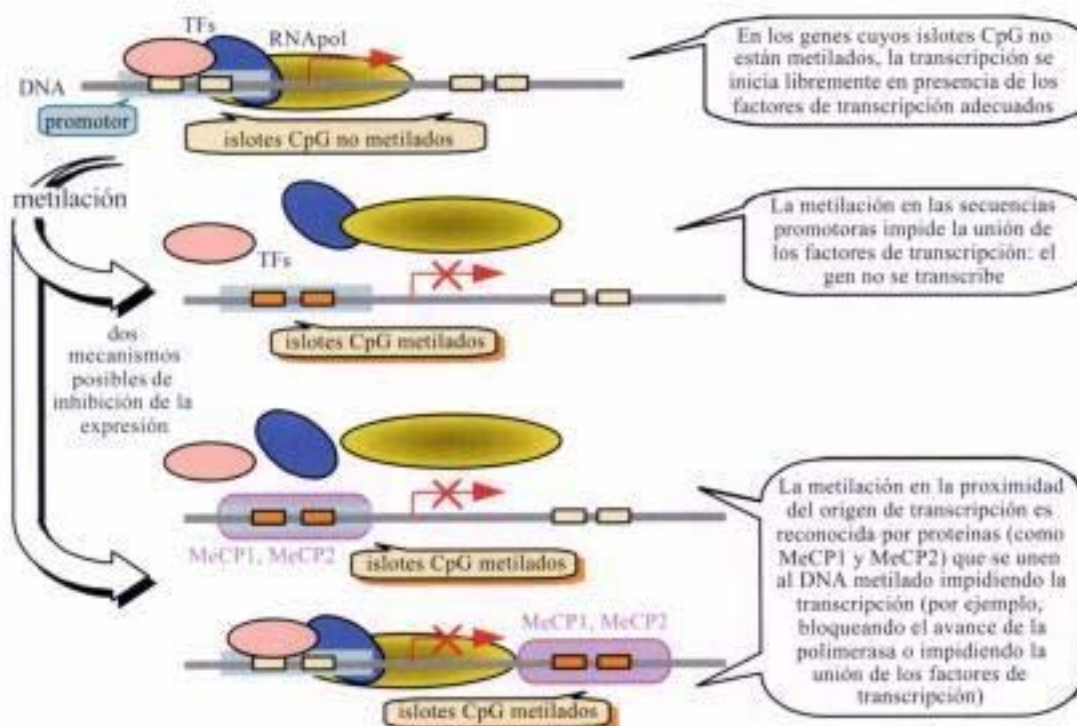
20.4 REGULACIÓN EPIGENÉTICA

Este segundo tipo de regulación de la expresión génica, que contribuye a explicar la diversidad morfológica y funcional de células con un mismo genoma, depende del llamado **componente epigenético** de la estructura del DNA, es decir, no consistente en su secuencia. Concretamente, se ejerce a través de la metilación específica de algunas bases del DNA. Aunque sin efecto regulador, se conoce el papel de la metilación del DNA propio en las bacterias, como protección frente a las enzimas de restricción de la misma célula (pág. 200). Corresponde ahora estudiar el papel de la metilación del DNA en células de vertebrados, para "silenciar" la transcripción de sus genes.

20.4.1 Metilación del DNA y control de la expresión

En el DNA de vertebrados una parte de los residuos de citosina (en humanos, el 3%) se encuentran metilados, principalmente en posición 5. Esta metilación aparece casi exclusivamente sobre la secuencia dinucleotídica CG, que además es especialmente abundante en las regiones promotoras de los genes, dando lugar a los denominados **islotos CpG**. Los genes que se transcriben activamente en ciertos tejidos poseen islotos CpG sin metilar (o islotos *hipometilados*), mientras que en los tejidos donde el gen no se expresa sí están metilados. Por tanto, la metilación de la citosina parece desempeñar un papel en la regulación de la expresión génica.

La detención de la transcripción de un gen como consecuencia de la presencia de grupos metilo en la molécula de DNA se puede explicar mediante dos mecanismos, no excluyentes:

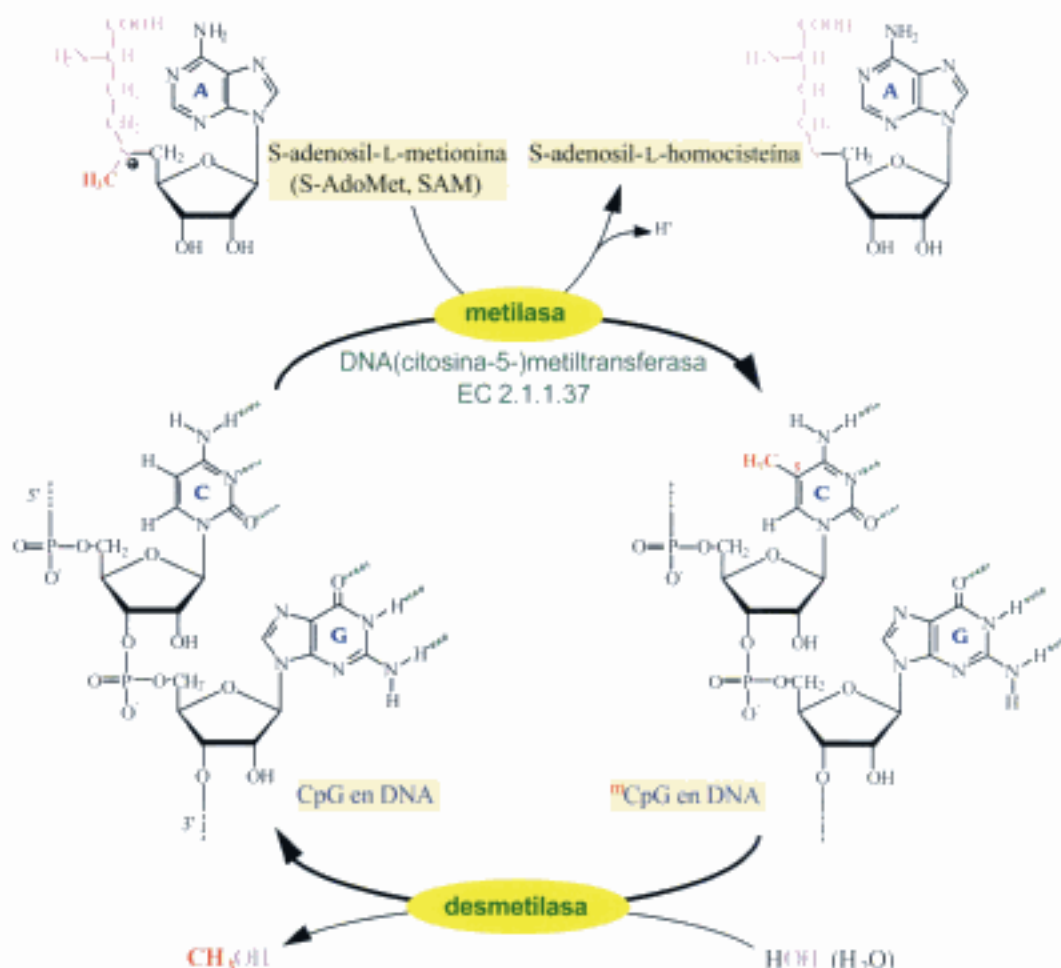


20.4.2 Patrón de metilación durante el desarrollo

Como manifestación de esta regulación por metilación, cada tejido presenta un patrón característico de metilación del DNA en sus células (componente epigenético). Éste se consigue y conserva durante el desarrollo embrionario, mediante **tres procesos sucesivos** de distribución de los grupos metilo:



De los procesos de metilación se encarga una **DNA metiltransferasa** (o **metilasa**), aunque aún no se sabe con certeza si la metilación *de novo* y la de mantenimiento dependen de una misma enzima o de dos diferentes. Se ha podido aislar esta enzima de muchos organismos, incluyendo los humanos, e incluso se han clonado sus genes. Actúa similarmente a las metilasas que forman parte de los sistemas de metilación-restricción en procariotas (pág. 200) y de la maduración de mRNA en eucariotas (pág. 278), empleando S-adenosilmetionina como coenzima donador de grupos metilo (pág. 22) y, en este caso, con especificidad por citidina adyacente a guanosa



La reacción opuesta, responsable del proceso de desmetilación global, la cataliza una **desmetilasa**, que reconoce las secuencias metiladas mCpG eliminando el grupo metilo de la citosina. Esta enzima se ha podido purificar a partir de células tumorales, en las que es especialmente abundante.

20.4.3 Metilación y cáncer

Estos procesos de metilación-desmetilación del DNA se han podido investigar gracias a la observación en células cancerosas de alteraciones en la metilación del DNA, posiblemente responsables de la modificación en el programa de expresión génica característica del proceso tumoral, aunque no su causa única. En esas células el DNA está globalmente **hipometilado**, pero al mismo tiempo algunos genes están **hipermetilados** (concretamente, los genes *supresores de tumores*, tema 30). Esta aparente contradicción es el resultado, por mecanismos complejos, de una elevada expresión de ambas enzimas, la metilasa y la desmetilasa. La hipometilación provocaría la expresión de genes que habitualmente están silenciados (incluido el de la desmetilasa), conduciendo posiblemente a algunos de los desajustes causantes del cáncer, mientras que la hipermetilación de los genes supresores de tumores bloquea su expresión, eliminándose su función natural, el freno de la proliferación celular, y conduciendo a la multiplicación anárquica de las células que genera el tumor (tema 30).

Tema 21

Maduración del RNA o procesamiento postranscripcional

21.1	Introducción	275
21.1.1	Matizaciones al concepto de gen	276
21.2	Características diferenciales de la maduración	276
21.3	Procesamiento del RNA mensajero	277
21.3.1	Modificación del extremo 5'	277
21.3.1.1	Formación de la caperuza o casquete 5'	278
21.3.1.2	Metilación de la caperuza 5'	278
21.3.2	Modificación del extremo 3'	279
21.3.2.1	Formación del extremo 3' y su relación con el final de la transcripción	279
21.3.2.2	Poliadenilación del extremo 3'	280
21.3.3	Eliminación de intrones y empalme de exones	280
21.3.3.1	Magnitud de los intrones	280
21.3.3.2	Secuencias que delimitan el intrón	280
21.3.3.3	Formación de cuerpos o complejos de empalme	280
21.3.3.4	Mecanismo de la reacción	281
21.3.3.5	Otros tipos de intrones	282
21.4	Procesamiento de los RNAs ribosómico y de transferencia	282
21.5	Maduración del RNA mitocondrial	284
21.6	Regulación postranscripcional y pretraducción de la expresión génica	284
21.6.1	Regulación de la maduración del RNA	285
21.6.2	Vida media y degradación del mRNA	286

21.1 INTRODUCCIÓN

La transcripción del DNA es un proceso que no difiere, en esencia, entre procariotas y eucariotas: las RNA polimerasas se unen a secuencias promotoras específicas, y se realiza la síntesis en dirección 5'→3' a lo largo del gen. Sin embargo, surgen importantes diferencias entre ambos tipos de células cuando se considera la formación del RNA funcional.

Se llama **transcrito primario o precursor del RNA** a la molécula de RNA que resulta directamente del proceso de transcripción. Todos los transcritos primarios tienen la misma secuencia que el fragmento correspondiente de la hebra no-molde (no transcrita) de DNA, salvo por la presencia de U en lugar de T. Existen transcritos primarios para los tres tipos de RNAs en eucariotas y para los tRNAs y rRNAs en procariotas. En contraste, el mRNA de procariotas es el único RNA que carece de un precursor, lo que permite su empleo inmediato para la traducción (que ocurre incluso antes de finalizar su síntesis por transcripción, pues no existe compartimentación subcelular, pág. 250).

Los precursores de los tres tipos de RNA formados en el núcleo eucariótico no pueden desempeñar directamente su función (en el caso del mRNA, utilizarse como molde para la traducción), sino que han de convertirse antes en RNA **maduro**, funcional, en un proceso llamado **modificación o procesamiento postranscripcional** o simplemente **maduración del RNA**. Ésta tiene lugar en el núcleo, y sólo cuando se ha completado pueden las moléculas de RNA transportarse al citoplasma, donde ejercen su función. Este transporte tiene lugar por paso de las moléculas de RNA maduro a través de los poros de la membrana nuclear (o **poros nucleares**).

El procesamiento posttranscripcional consiste, de modo general, en una o varias de las siguientes reacciones (véase también pág. 283):

- Modificación covalente de nucleósidos
- Adición de nucleótidos a los extremos
- Escisión de nucleótidos de los extremos
- División de la molécula en varios fragmentos
- Proceso de ajuste, eliminación de intrones, o corte de intrones y empalme de exones

Como se verá más adelante, la maduración confiere estabilidad al RNA, básicamente por una mayor resistencia frente a las nucleasas y por un plegamiento tridimensional compacto (en el caso de tRNAs y rRNAs). Igualmente, facilita su reconocimiento por otros componentes celulares, que mejora la funcionalidad para la traducción.

21.1.1 Matizaciones al concepto de gen

La maduración posttranscripcional es uno de los procesos que dan lugar a situaciones en las que la definición de **gen** debe matizarse o forzarse ligeramente (pág. 246). Concretamente, pueden citarse dos situaciones:

- Un gen se transcribe dando un pre-RNA, que es procesado por escisión generando varias moléculas de RNA. En este caso, el concepto de gen se puede aplicar al gen "global" de partida, puesto que es una unidad de transcripción, con sus regiones estructural y reguladora. Sin embargo, también es aplicable a cada una de las regiones "parciales" de DNA que codifican uno de los RNAs finales, pues son éstos los productos génicos funcionales. Ejemplos de esta situación los tenemos en los precursores de RNA ribosómico, tanto procarióticos como eucarióticos (pág. 283) cuyo "gen global" da lugar a varios rRNAs maduros, y también en el DNA mitocondrial, cada uno de cuyos transcritos primarios se procesa dando una combinación de rRNAs, tRNAs y mRNAs (pág. 284).
- El transcrito primario obtenido de una región de DNA puede sufrir su maduración por dos rutas o variantes distintas, dando lugar por ello a dos productos génicos diferentes. Este es el caso del **ajuste alternativo** (pág. 285), donde se pueden emplear dos secuencias de corte y empalme para la eliminación de intrones (pág. 280) dependiendo, por ejemplo, del tejido en el que se exprese el gen o de la llegada a la célula de diversas señales.

21.2 CARACTERÍSTICAS DIFERENCIALES DE LA MADURACIÓN

Las diferencias entre procariotas y eucariotas relativas a la maduración del RNA se centran en el RNA mensajero; los tRNAs y rRNAs sufren una maduración similar en ambos tipos de organismos. Puesto que, además, afecta a la síntesis de proteínas, nos centraremos, por tanto, en el análisis de la maduración del mRNA.

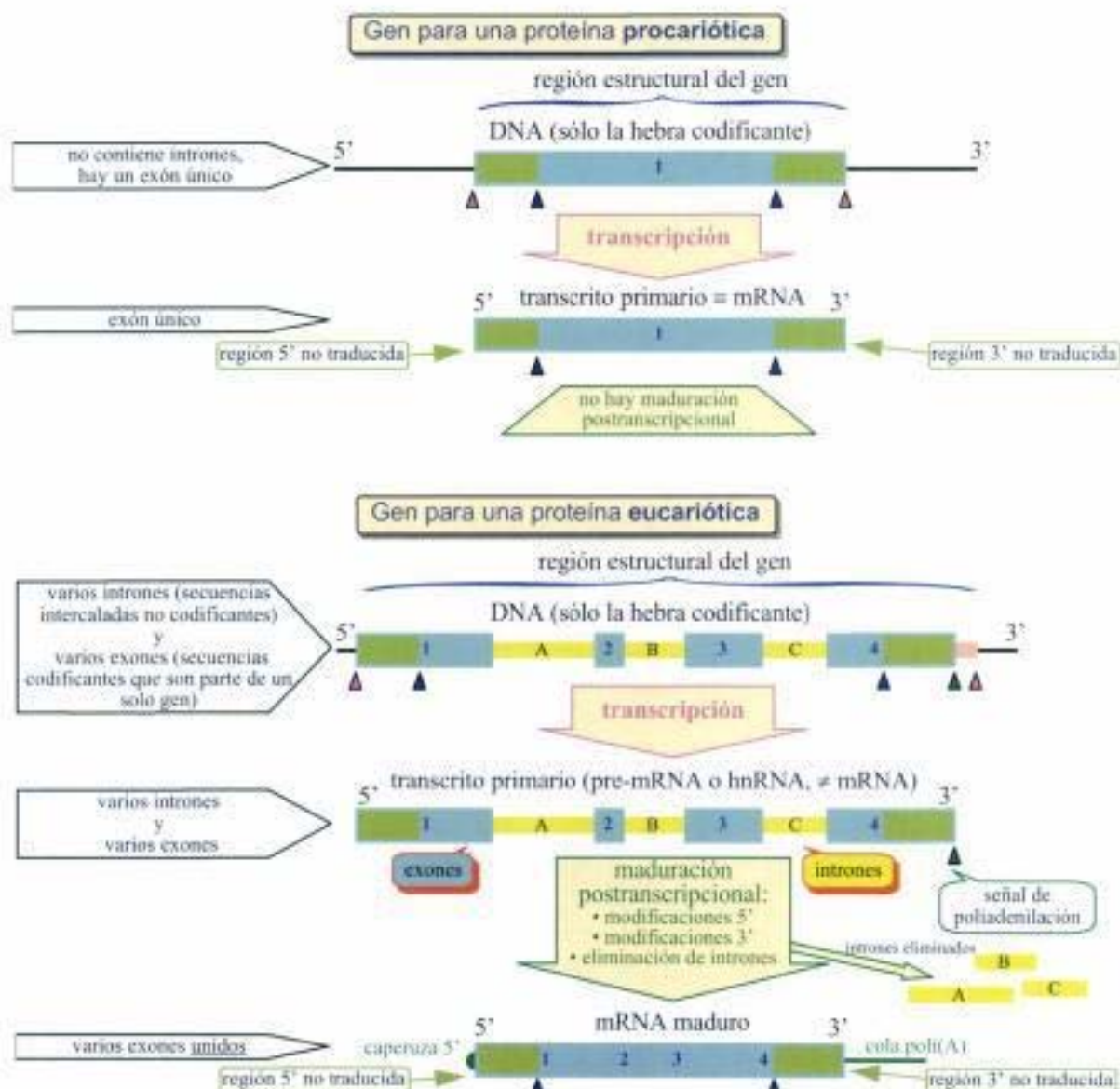
En un gen proteico **procariótico** todo el DNA es codificante (excluyendo la región reguladora, que en este tema no nos interesa considerar), de modo que al transcribirse origina un transcrito primario que es ya el mRNA funcional, cuya secuencia corresponde exactamente a la del DNA. Es decir, el mRNA de procariotas es el único RNA que no requiere maduración. Por el contrario, la transcripción de un gen proteico **eucariótico** genera un pre-mRNA que debe sufrir maduración para transformarse en una molécula funcional de mRNA. El conjunto de los diferentes pre-mRNAs se llama **RNA nuclear heterogéneo (hnRNA)**, debido a que constituye una población de moléculas de tamaño muy variado.

Dentro de las diferencias entre pre-mRNA y mRNA maduro, la más llamativa es que la secuencia del mRNA no corresponde completamente a la del pre-mRNA, ni a la del DNA del que se ha transcrito (se dice que "se ha perdido la colinealidad entre DNA y RNA"). Esto se debe a que grandes segmentos internos del gen desaparecen durante la maduración. Estas *regiones del DNA que se eliminan del pre-mRNA en su maduración* se denominan **intrones**. Aunque son DNA no codificante (pág. 108), deben contemplarse como parte del gen (pág. 245).

Se pueden definir también los intrones como *regiones no codificantes situadas en el interior de un gen*, lo que da lugar al sinónimo **secuencias intercaladas (IVS, intervening sequences)**. No son intrones aquellas secuencias que forman los extremos 5' y 3' del mRNA maduro, aunque se trate igualmente de DNA no codificante; por ello, una tercera definición de intrón como *región que se transcribe pero no se traduce* no es estrictamente correcta.

Aunque los intrones son característicos de los genes eucarióticos que codifican proteínas, también están presentes en algunos genes de tRNA y rRNA, en especial eucarióticos. En éstos el **intrón** se puede definir únicamente como la región de DNA (o de su producto transcripcional, el pre-RNA) que se elimina durante la maduración.

Finalmente, las regiones que flanquean a los intrones y que se conservan en el RNA maduro se denominan a su vez **exones**. En general, son la parte codificante del gen, pero el primer exón y el último también incluyen, respectivamente, las regiones 5' y 3' no traducidas, que no son codificantes.



Por tanto, el DNA de los genes **procarióticos** de proteínas se puede considerar formado por un exón único, sin intrones, que al transcribirse origina un transcrito primario también sin intrones. Sin embargo, casi todo el DNA génico de **eucariotas** contiene varios intrones y exones, y se transcribe a precursores de RNA con los mismos intrones y exones. Como parte del proceso de modificación postranscripcional se liberan los intrones y se unen los exones.

La formación del mRNA eucariótico maduro a partir de su correspondiente pre-RNA constituye un modelo ideal para el estudio del procesamiento postranscripcional, pues sufre todos los tipos de transformación implicados en dicho proceso: adición de nucleótidos al extremo 5', metilación posterior, alargamiento de la cadena en el extremo 3' y eliminación de intrones con unión de exones ("corte y empalme"). Una vez completada esta maduración, también se produce el paso del mRNA maduro a través de los poros de la membrana nuclear hacia el citoplasma, lugar donde se empleará para la síntesis proteica. Se estudiarán con detenimiento cada uno de estos pasos.

21.3 PROCESAMIENTO DEL RNA MENSAJERO

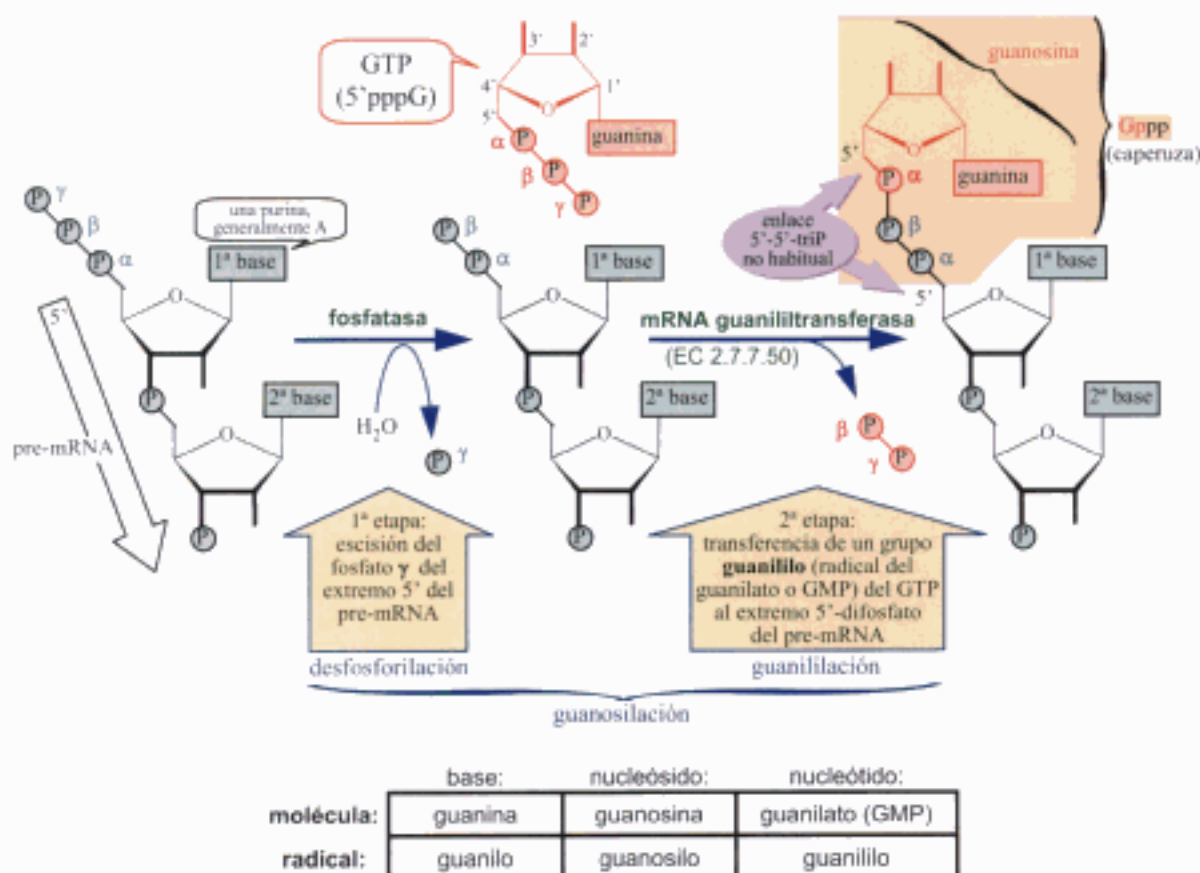
La maduración del mRNA puede describirse mediante tres tipos de modificaciones, que pueden transcurrir de forma sucesiva o simultánea, todas ellas realizadas en el núcleo.

21.3.1 Modificación del extremo 5'

El extremo libre 5'-PPP del mRNA eucariótico comienza a modificarse muy pronto, poco después de iniciada la transcripción. Se trata de una modificación covalente que, globalmente, implica tres tipos de enzimas (fosfatasa, guaniltransferasa y metilasas) y que conduce a la incorporación singular de un nucleótido protector sobre el NTP en posición 5' terminal del RNA (normalmente purínico, ATP, el mismo que actuó como cebador en la transcripción).

21.3.1.1 Formación de la caperuza o casquete 5'

Se realiza en dos etapas: una **desfosforilación** previa y una **guanililación** a costa de GTP.



Aunque en la segunda reacción se transfiere un residuo guanililo, el resultado conjunto de las dos reacciones es la incorporación al mRNA de tan sólo un residuo de guanosina, es decir, una **guanosilación**; se dice que el mRNA ha sido guanilado en 5'. La estructura terminal resultante se conoce como **caperuza de guanina**, **caperuza 5'** o **casquete 5'** (en inglés, *cap*). Se trata de una estructura singular, que no se encuentra en ninguna otra parte de nucleótidos ni ácidos nucleicos, con respecto a los siguientes criterios:

- Unión de dos nucleósidos a través de un enlace trifosfato.
- Un nucleósido (guanosina) unido en sentido opuesto al del resto de la cadena polinucleotídica (enlace trifosfato entre dos posiciones 5').

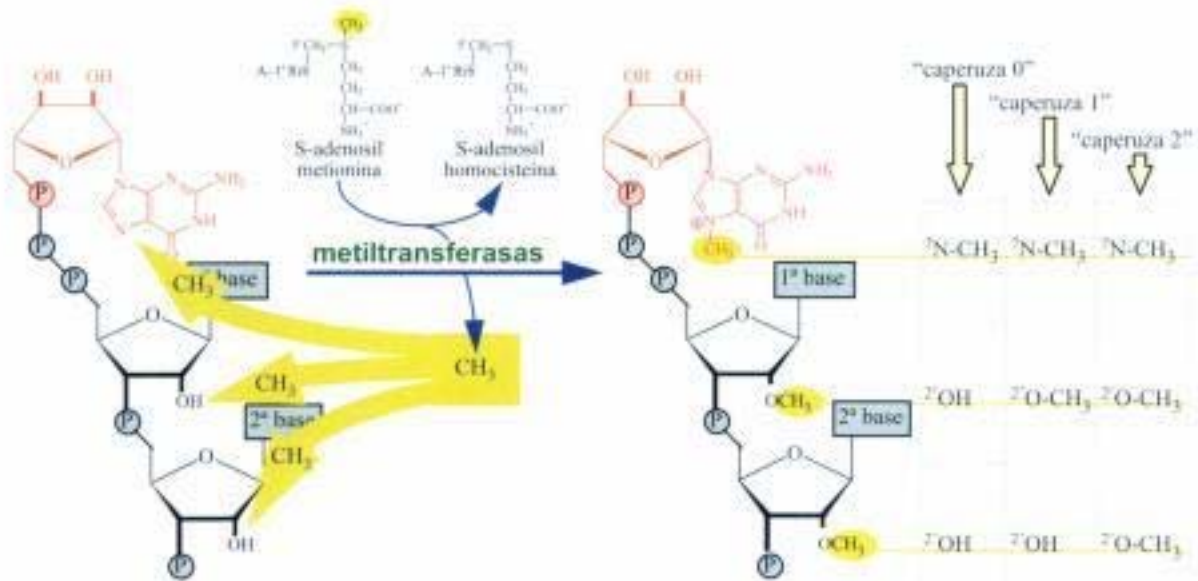
Esta caperuza ejerce un efecto protector de la molécula frente a las exonucleasas (pues éstas sólo actúan sobre enlaces 3'-5'-fosfodiéster), marca el pre-mRNA para su empleo posterior como sustrato en otras reacciones de procesamiento en el núcleo y sirve como punto de unión del mRNA a los ribosomas al iniciar la síntesis proteica.

21.3.1.2 Metilación de la caperuza 5'

Los tres primeros nucleótidos de la estructura 5'-guanosina-trifosfato-mRNA son entonces modificados por metiltransferasas (metilasas), que emplean S-adenosil-L-metionina (pág. 22) como coenzima donadora del grupo metilo (reacción similar a la de metilación de citosinas que regula la transcripción, pág. 272).

La metilación puede tener lugar sobre varias posiciones:

- El N-7 de la guanosina terminal, para dar la **caperuza de tipo 0**, o caperuza de 7-metilguanosina-trifosfato (*cap 0*). La reacción la cataliza la mRNA(guanina-N7-)metiltransferasa, EC 2.1.1.56.
- Además del anterior, el 2'-OH de la ribosa adyacente (1^{er} nucleótido original del pre-mRNA; 2^o de la secuencia actual), dando la **caperuza de tipo 1** (*cap 1*). Esta reacción la cataliza una metilasa diferente, la mRNA(nucleósido-2'-O)metiltransferasa, EC 2.1.1.57.
- Además de las dos anteriores, puede metilarse también el 2'-OH de la ribosa siguiente (2^o nucleótido original; 3^o de la secuencia actual), para dar la **caperuza de tipo 2** (*cap 2*).

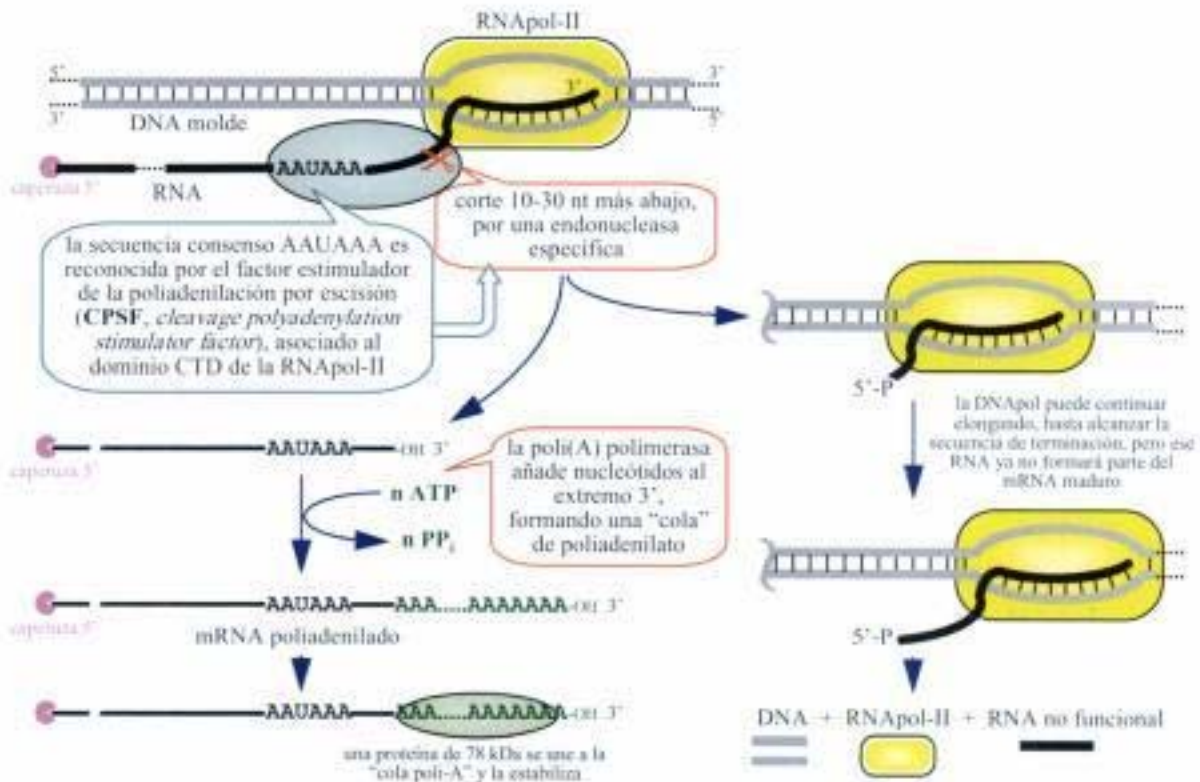


Se sabe que las metilasas y las guaniltransferasas se asocian al dominio CTD fosforilado de la RNAPol-II (pág. 252) y llevan a cabo la reacción de procesamiento del RNA antes de que el transcrito haya alcanzado un tamaño de unos 30 nucleótidos. Parece, por tanto, posible que la formación de la caperuza participe en el cambio de iniciación de la transcripción a elongación.

21.3.2 Modificación del extremo 3'

21.3.2.1 Formación del extremo 3' y su relación con el final de la transcripción

Otra peculiaridad del mRNA eucariótico es que el extremo 3' del mRNA no corresponde a la posición donde se termina la transcripción (que se produce gracias a la interacción de las RNAPol con proteínas de pausa o terminadoras pág. 256), sino que deriva de la escisión de la molécula de pre-mRNA en una posición corriente arriba de su extremo 3'. Como consecuencia, el mRNA es más corto en su extremo 3' que su transcrito primario.



21.3.2.2 Poliadenilación del extremo 3'

Sobre el extremo 3'-OH resultante del corte por la endonucleasa actúa una RNA polimerasa especial, que no usa molde y sólo acepta como sustrato al ATP, llamada **poli(A) polimerasa** (EC 2.7.7.19). Como consecuencia, se añade al RNA un gran número de residuos de adenilato (AMP), formando una "cola" de poliadenilato o **cola de poli(A)**, de longitud típica entre 40 y 250 nt (como excepción, los mRNAs de histonas carecen de esta cola). Esta estructura, que permanece intacta hasta la formación final del mRNA, puede participar en procesos de tan gran significado como el transporte del mRNA al citoplasma y la determinación, en parte, de la estabilidad y vida media del mRNA (pág. 286).

La presencia de la cola de poli(A) únicamente en los mRNAs eucarióticos se aprovecha con fines experimentales, para la purificación del conjunto de mRNAs de una célula, empleando técnicas de afinidad con soportes que contienen cadenas oligonucleotídicas de poli(dT), poli(T) o poli(U), complementarias a la cola de poli(A) (pág. 136).

21.3.3 Eliminación de intrones y empalme de exones

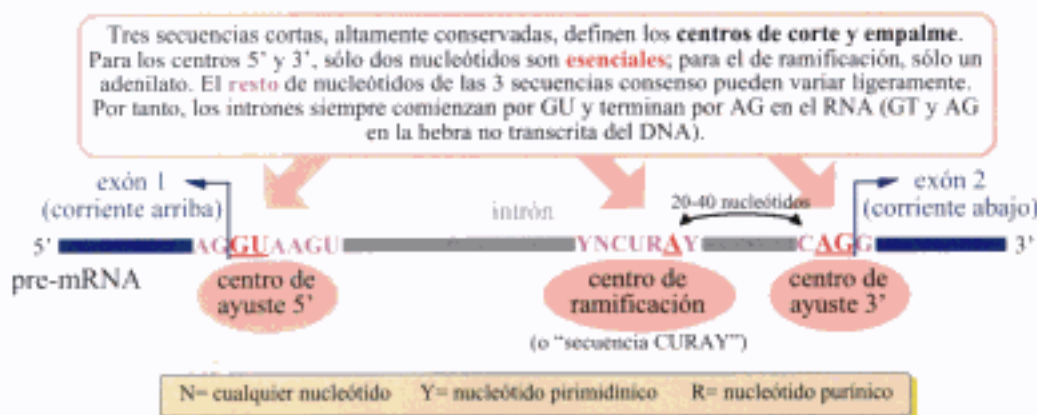
La etapa esencial de la maduración postranscripcional del transcrito primario hasta dar el mRNA eucariótico consiste en la eliminación de los intrones y la unión o empalme de los exones. Este proceso se conoce como **corte y pegado, corte y empalme** o **ajuste** (en inglés, *splicing*).

21.3.3.1 Magnitud de los intrones

Debe recordarse que los genes de eucariotas, especialmente en los vertebrados, se encuentran casi siempre interrumpidos por **intrones**, que separan las regiones codificantes o **exones** (pág. 276). Como excepción, las histonas son uno de los escasos ejemplos de genes que carecen completamente de intrones. Lo más frecuente es que la longitud total de los intrones supere ampliamente a la suma de los exones, por lo que el tamaño del gen depende fuertemente del número y longitud de sus intrones. Aunque las cifras son variables, se habla de hasta 40 intrones por gen y de una longitud de hasta 20 kb por intrón. Los exones son de tamaño mucho más pequeño, típicamente inferior a 1 kb.

21.3.3.2 Secuencias que delimitan el intrón

La eliminación de un intrón está determinada por su secuencia, y no por su tamaño. Sólo intervienen dos secuencias específicas, que definen los extremos del intrón, y otra de su interior, conocidas en su conjunto como **símbolos o centros de corte y empalme** o **de ajuste** (centro 5', centro 3' y centro de ramificación). Todos los sitios de corte y empalme parecen ser equivalentes, mientras que la secuencia del resto del intrón no tiene trascendencia sobre el proceso.



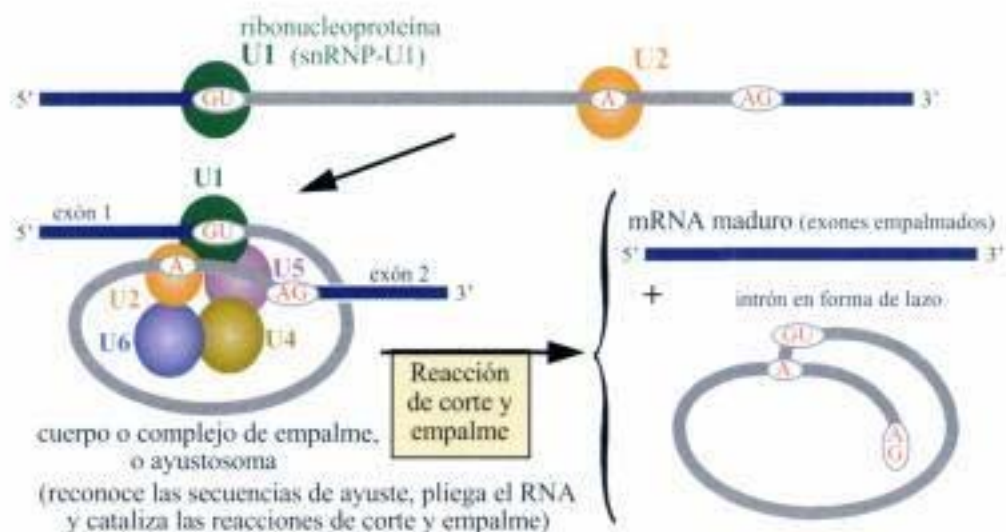
Un cambio en el sitio de corte (por ejemplo, por mutación de uno de los nucleótidos consenso esenciales) daría lugar a un ajuste erróneo: se utilizaría para el corte y empalme la siguiente secuencia consenso, si la hubiera, con lo que el intrón o parte de él quedaría incluido en el mRNA. Esto conduciría a la síntesis de un polipéptido alterado, con una inserción grande en su secuencia de aminoácidos, o incluso con cambio de toda la secuencia a partir del punto de alteración debido a un desplazamiento del marco de lectura (pág. 350).

21.3.3.3 Formación de cuerpos o complejos de empalme

El corte de intrones y empalme de exones está mediado por una asociación macromolecular denominada cuerpo o complejo de empalme, o **ajustoso** (*spliceosome*). Éste se forma por asociación del pre-mRNA (alrededor de los centros de ajuste) con varias **ribonucleoproteínas nucleares pequeñas** (**snRNP**, denominadas U1, U2, U4, U5 y U6), constituidas a su vez por proteínas y **RNA nucleares pequeños** (**snRNA**, del mismo nombre). El tamaño del

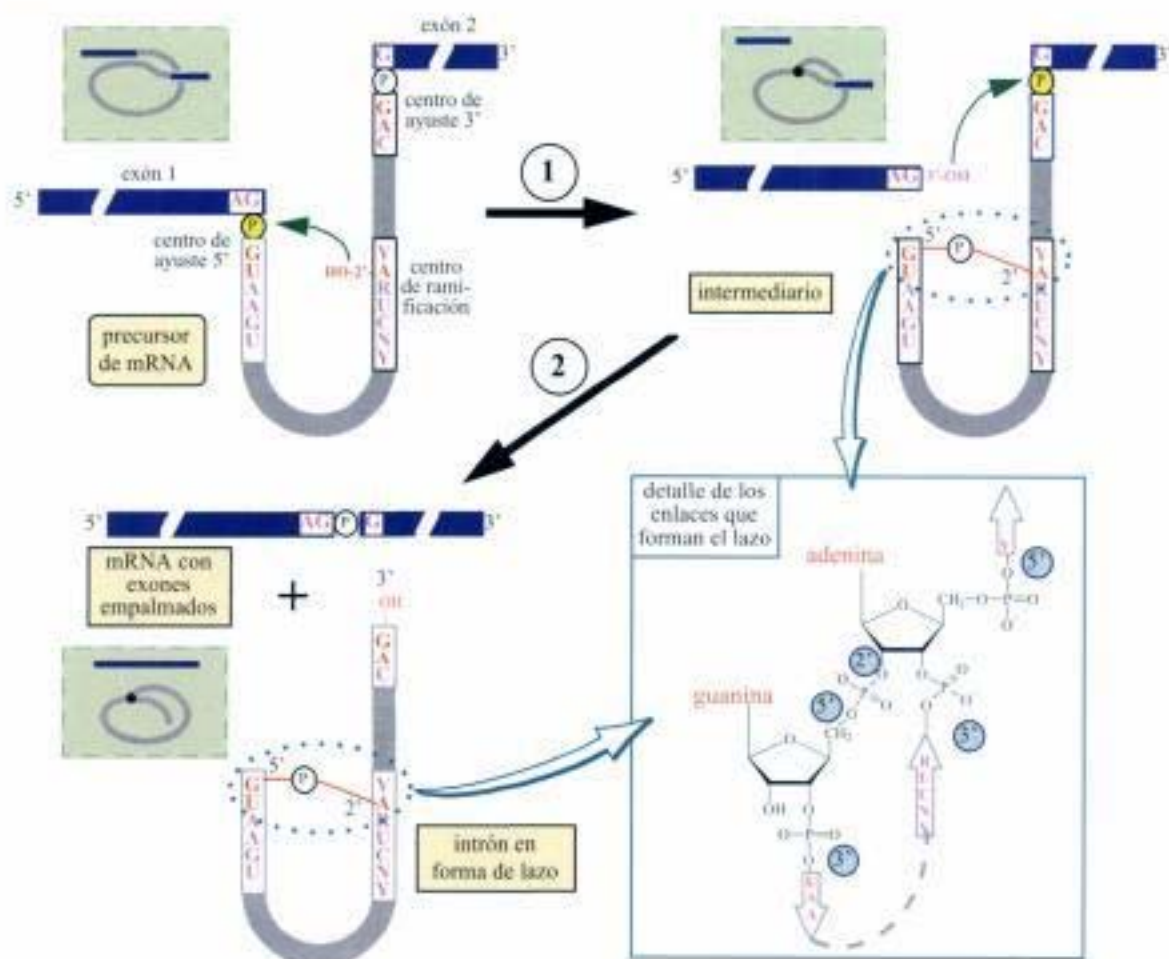
ayustosoma es de unos 3.000 kDa, con un coeficiente de sedimentación de 60S, similar a la subunidad mayor del ribosoma eucariótico; en total, está formado por alrededor de 45 polipéptidos y 5.000 nt de RNA.

Este complejo de corte y empalme es necesario tanto para el reconocimiento de los puntos de escisión como para la catálisis.



21.3.3.4 Mecanismo de la reacción

La reacción de corte y empalme supone, aparentemente, la ruptura de dos enlaces fosfodiéster (los que separan el intrón de ambos exones) y la formación de un nuevo enlace fosfodiéster entre los dos exones. Sin embargo, el mecanismo catalizado por el ayustosoma no transcurre de esa forma, no hay acción de nucleasas y ligasa, sino dos reacciones de **transesterificación**:



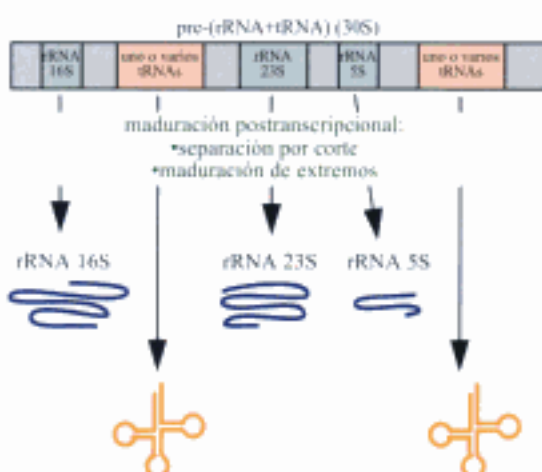
En **eucariotas**, el proceso es similar, con un transcrito primario que genera 3 de los 4 rRNAs y con transcritos primarios separados para el otro rRNA y para cada tRNA. También en algunos casos hay eliminación de intrones.

En el caso de tRNAs se produce, además de un recorte de los extremos 3' y 5', la adición, a las moléculas de tRNA que no la posean, de la secuencia -CCA en el extremo 3'-OH. Los tRNAs, especialmente, sufren numerosas reacciones de modificación química que determinan que en las moléculas maduras abundan los nucleósidos infrecuentes (págs. 13, 21, 72 y 74), diferentes a los cuatro que participan en la síntesis del RNA (A, G, U, C). Estas modificaciones contribuyen a hacer más resistentes los tRNAs a los procesos de degradación celular, aumentando, por tanto, su estabilidad metabólica y su vida media.

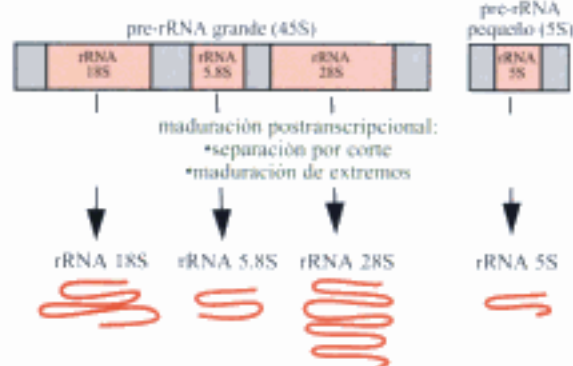
Resumen de las modificaciones postranscripcionales experimentadas por los distintos tipos de RNA

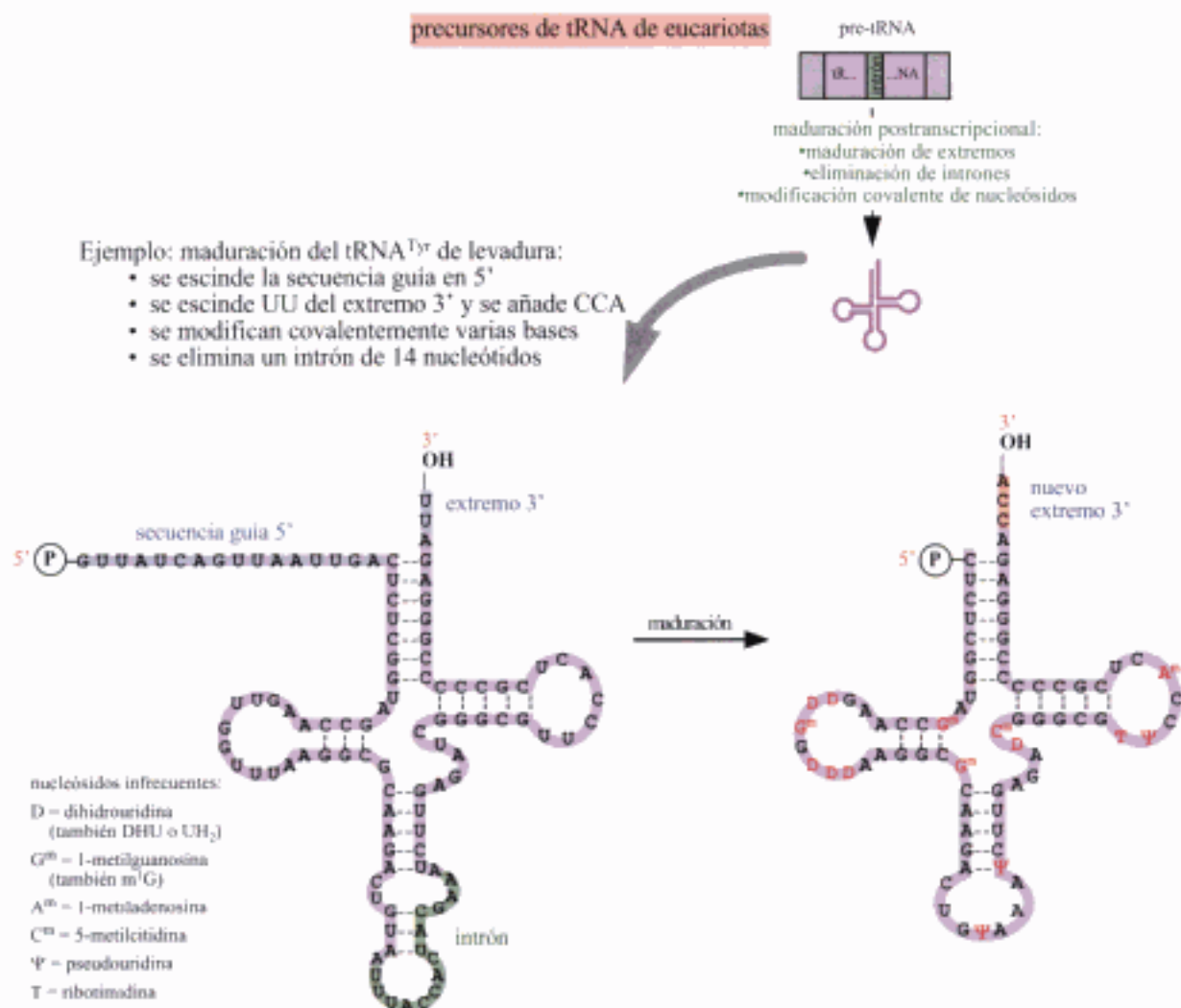
	precursor	corte por endonucleasas	corte de intrones y empalme de exones	corte por exonucleasas	adición en 3'	modificación en 5'	modificación de nucleósidos
PROCARIOTAS							
mRNA	----- (no sufren procesamiento postranscripcional) -----						
rRNA	de 30S, da lugar a los tres rRNAs (16S, 23S, 5S) y algún tRNA	a ambos lados (RNasa III)	en algunos organismos (autoempalme)	procesamiento en ambos extremos			algunas bases metiladas
tRNA	da lugar a varios tRNAs y en algunos casos también a rRNAs	a ambos lados (RNasa P y RNasa III)		extremo 3' recortado por RNasa D			muchas bases modificadas, gran diversidad
EUCARIOTAS							
mRNA	pre-mRNA (da lugar a un solo mRNA)	sólo en el extremo 3'	sí (con autoempalme en mRNA mitocondrial)		cola de poli(A)	caperuza 5'	en casos puntuales (ej.: desaminación de C→U)
rRNA	de 35 a 47S, da lugar a tres rRNAs (18S, 5.8S, 28S); otro precursor da el rRNA 5S	(similar a procariotas)	en algunos unicelulares (autoempalme)	(similar a procariotas)			metilación de ribosas
tRNA	da lugar a un solo tRNA		en algunos; distinto mecanismo	en ambos extremos	adición de CCA (tRNA-nucleotidil-transferasa)		muchas bases modificadas, menor diversidad

precursor mixto de rRNAs y tRNAs de procariotas



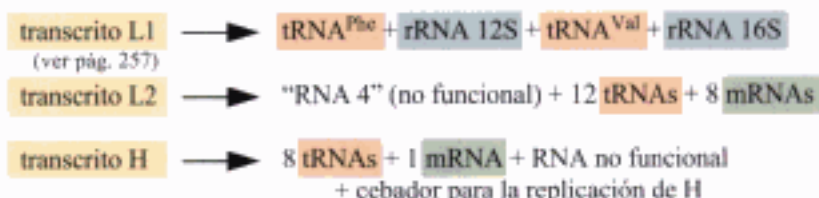
precursores de rRNA de eucariotas





21.5 MADURACIÓN DEL RNA MITOCONDRIAL

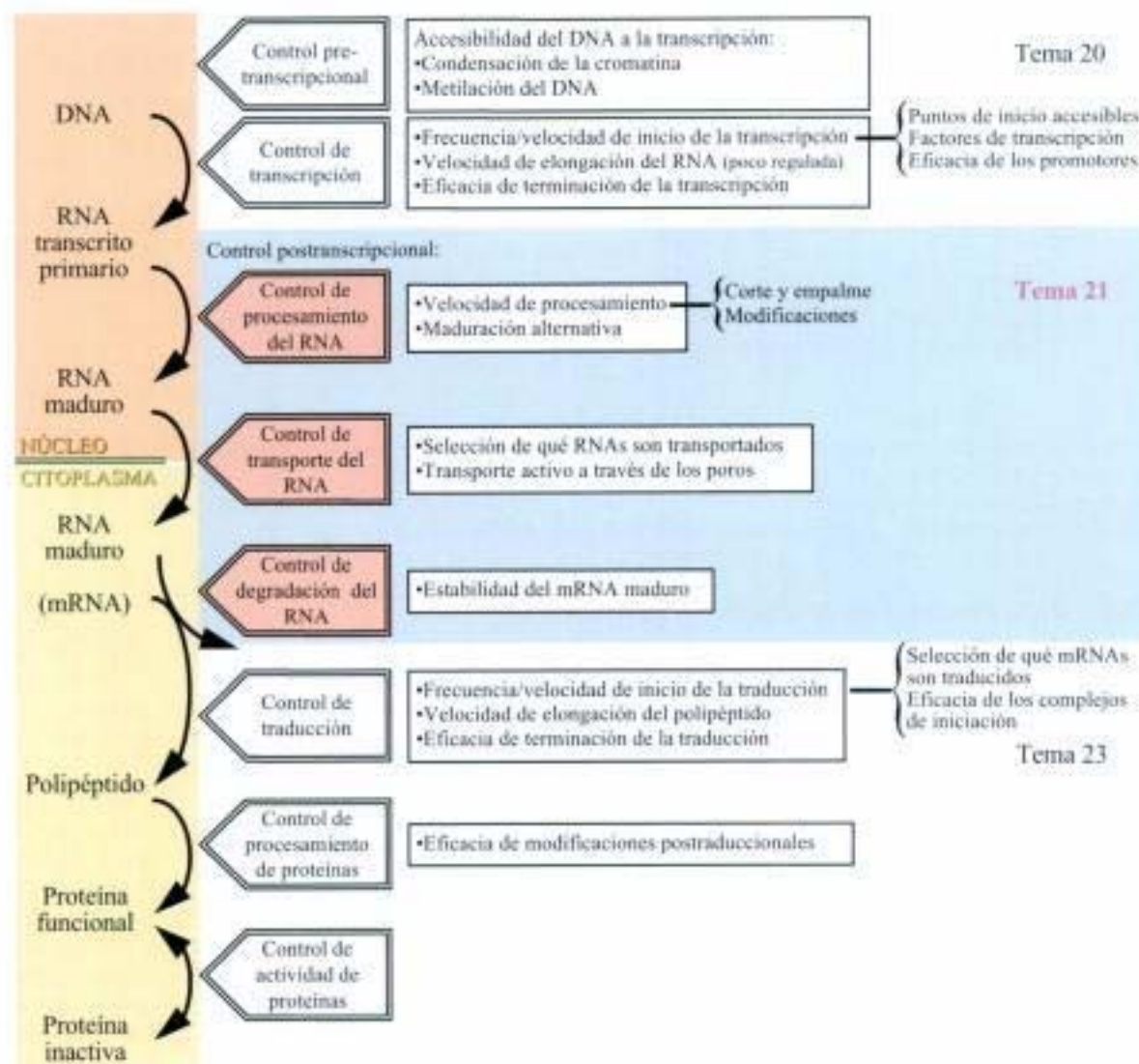
La transcripción de la molécula de DNA mitocondrial humano (pág. 257) produce 3 RNAs transcritos primarios de gran tamaño, cuyo procesamiento da lugar a todos los mRNAs, tRNAs y rRNAs. Esta maduración tiene lugar principalmente por escisión del pre-RNA y, en el caso de los mRNAs, también por poliadenilación (pero no hay formación de caperuza 5').



21.6 REGULACIÓN POSTRANSKRIPCIONAL Y PRETRADUCCIONAL DE LA EXPRESIÓN GÉNICA

A pesar de que la mayor parte del control de la expresión génica se ejerce durante la etapa de inicio de la transcripción, a través de los factores de transcripción (pág. 264), existen mecanismos de regulación en todas las etapas de la expresión (pág. 262). En lo que respecta a los procesos posteriores a la transcripción pero previos a la traducción, existen ejemplos de control tanto en la maduración del RNA como en su transporte al citoplasma y también a través de la regulación de la vida media del mRNA en el citoplasma (velocidad de su degradación). Por razones de simplicidad, se desarrolla únicamente este último aspecto.

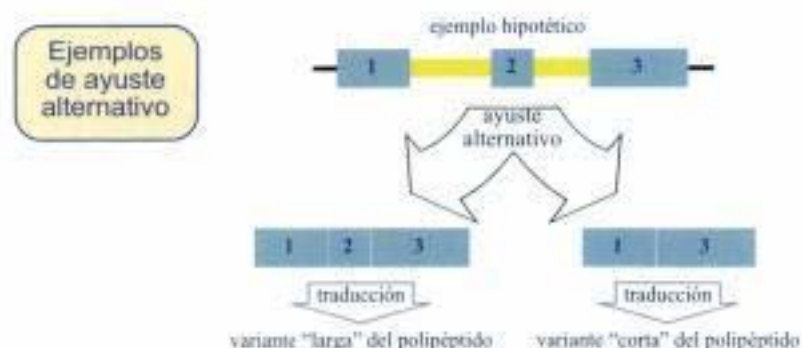
NIVELES DE REGULACIÓN DE LA EXPRESIÓN GÉNICA EN EUCARIOTAS



21.6.1 Regulación de la maduración del RNA

Las distintas etapas de procesamiento de los pre-RNAs, previamente estudiadas, son lógicamente susceptibles de regulación, lo que afectará a la tasa de expresión de los genes correspondientes. Esa regulación no sólo tiene lugar en cuanto a la velocidad o eficacia de las reacciones de maduración, sino que en algunos casos un mismo pre-RNA madura según procesos distintos dependiendo de la célula considerada, formando mRNAs adaptados a las necesidades proteicas particulares de un tejido o de una situación fisiológica concreta. La implicación de este tipo de variabilidad en el control del producto final es enorme y, por ahora, poco conocida; a este nivel, tendrá mucho que aportar en el futuro la proteómica (pág. 217).

Un ejemplo característico de esta forma de regulación es el **ajuste alternativo**, que comúnmente desempeña un papel en la diferenciación celular y en el control del desarrollo (por ejemplo, en la determinación del sexo en *Drosophila*).



Tema 22

El código genético

22.1 Antecedentes y propiedades generales del código	287
22.1.1 Los codones del código son tripletes	288
22.1.2 Los tripletes no se solapan	288
22.1.3 Existen tres marcos o pautas de lectura	288
22.1.4 Papel adaptador del tRNA	289
22.2 Asignación de codones a aminoácidos concretos	289
22.2.1 Traducción de polinucleótidos sintéticos	289
22.2.1.1 Ensayo con homopolímeros	289
22.2.1.2 Ensayo con secuencias repetidas de dos nucleótidos	291
22.2.1.3 Ensayo con secuencias repetidas de tres nucleótidos	291
22.2.1.4 Ensayos con secuencias repetidas de cuatro nucleótidos	292
22.2.2 Evidencia <i>in vivo</i> del código genético	292
22.3 Modelos de representación	293
22.3.1 Formato en tabla	293
22.3.2 Formato circular	294
22.4 Características específicas	294
22.4.1 Codón de inicio	294
22.4.2 Codones de terminación	294
22.4.3 Degeneración	295
22.4.3.1 Hipótesis del balanceo	296
22.4.3.2 Consideraciones prácticas	297
22.4.4 Universalidad	298

El código genético es el conjunto de pautas que rigen la transferencia de la información contenida en el mRNA (y en definitiva en el DNA) para la síntesis de las proteínas. Describe, tanto en procariotas como en eucariotas, la correlación entre dos lenguajes informativos: la secuencia de nucleótidos en el mRNA (el *lenguaje con cuatro letras* de los ácidos nucleicos) y la secuencia de aminoácidos en el polipéptido (*lenguaje con veinte letras* de las proteínas). Más concretamente, el código genético establece la correspondencia entre los 64 posibles tripletes de bases del mRNA y los 20 aminoácidos de las proteínas. (**Nota:** Es frecuente, especialmente en el código genético y la traducción, utilizar la palabra "base" como sinónimo de "nucleótido", en lo relativo a la secuencia del mRNA).

22.1 ANTECEDENTES Y PROPIEDADES GENERALES DEL CÓDIGO

Hoy día es relativamente fácil conocer la correspondencia entre aminoácidos de una proteína y tripletes de nucleótidos del mRNA que los codifican, mediante simple comparación de la secuencia de ambos. Sin embargo, fueron muchas las aproximaciones experimentales realizadas, a partir de los años 50, que contribuyeron a descifrar el código genético, a la vez que a entender la transcripción y la traducción. Globalmente considerados, estos estudios han supuesto uno de los mayores retos en la historia de la Bioquímica y la Genética.

La información previa en la que se basa el conocimiento del código genético es de dos tipos. Por un lado, lo referente al tRNA que, a pesar de no ser un elemento propiamente dicho del código genético, su intervención es esencial para entender la correspondencia entre mRNA y proteína. Por otro lado, el **ribosoma** (págs. 79-81), orgánulo citoplasmático donde se realiza la síntesis proteica, según se demostró empleando aminoácidos marcados con isótopos radiactivos.

Se describen a continuación las características generales y particulares del código genético, basadas en algunos de los experimentos más representativos que han contribuido a descifrar el código genético y su relación con la síntesis proteica

22.1.1 Los codones del código son tripletes

La "hipótesis de los codones" se demostró al observar que cambios en la secuencia de un mRNA conducían a alteraciones en la secuencia de aminoácidos de la proteína sintetizada (en el bacteriófago T4). El número mínimo de nucleótidos necesario para "codificar un aminoácido" (tamaño de las unidades de codificación o relación de codificación) se deduce por un simple cálculo matemático:

unidad de información	nº de combinaciones posibles	
1 nucleótido	4	insuficiente para codificar
2 nucleótidos	$4^2 = 16$	20 aminoácidos
3 nucleótidos	$4^3 = 64$	más que suficiente
La unidad de codificación es el triplete de nucleótidos		

Por tanto, para poder codificar los 20 aminoácidos naturales se necesita que cada aminoácido venga especificado por una secuencia de 3 nucleótidos, llamada **tripleto** o **codón**. El tripleto es la característica que define la naturaleza general del código. Obviamente, las 64 posibles combinaciones exceden las necesidades para 20 aminoácidos; más adelante se analizará cómo se acomoda ese "exceso de información" (pág. 295).

22.1.2 Los tripletes no se solapan

Una de las propuestas iniciales para la distribución de los codones en el mRNA fue que tripletes contiguos compartieran nucleótidos.



En la hipótesis solapante, la secuencia del primer triplete restringe las posibles secuencias del segundo, éstas a su vez las del tercero, y así sucesivamente, por lo que la secuencia de aminoácidos de las proteínas estaría condicionada, no podría ser cualquiera. Bajo esta hipótesis la alteración de una base afecta a varios codones, lo que produciría alteraciones en más de un aminoácido. Puesto que estas previsiones no se cumplen en la realidad, se pudo demostrar que la hipótesis solapante es incorrecta, los tripletes nunca comparten nucleótidos, no se solapan. Esto es así para el código genético de todos los seres vivos.

22.1.3 Existen tres marcos o pautas de lectura

El hecho de que los codones contiguos no se solapan no implica que exista señal de separación entre ellos (es decir, no existen "comas", "puntos" o interrupciones; tal separación sólo se hace notar sobre el papel, para facilitar la lectura e interpretación de la secuencia). Como consecuencia, en principio la lectura de un mRNA en la traducción se podría iniciar en cualquier nucleótido, lo que definiría a éste como el primer nucleótido del codón de inicio de la traducción. Esto significa la existencia potencial de tres mensajes en cualquier mRNA, dependiendo de en qué nucleótido se inicie la lectura de los codones; se habla de tres **marcos o pautas de lectura**. La traducción del mRNA siguiendo cada uno de ellos formaría tres polipéptidos distintos.

3 formas de "leer" la secuencia del mRNA

3 marcos de lectura

1º: ABC, DEF, GHI, JKL, ...

2º: BCD, EFG, HIJ, KLM, ...

3º: CDE, FGH, IJK, LMN, ...

(el 4º ya coincide con el primero, el mensaje es el mismo aunque con una aminoácido menos)



Sin embargo, en la traducción (tanto *in vivo* como *in vitro*) no se manifiesta esta aparente ambigüedad (por ejemplo, al traducir el mRNA correspondiente a un gen). Una vez iniciada la traducción según un marco de lectura, se puede decir que los otros dos carecen de información. Sin embargo, uno de los otros marcos puede representar un mensaje válido para otro gen distinto, caso de los genes solapantes (pág. 249).

El marco de lectura puede cambiar debido al desplazamiento en 1 ó 2 posiciones del punto de inicio o, durante la traducción, por salto accidental de un nucleótido mientras el mRNA es recorrido por el ribosoma. La consecuencia del cambio de marco es que a partir de ese punto la secuencia de la proteína nueva es distinta a la original. Éste es el mecanismo utilizado a veces para formar dos o más proteínas relacionadas a partir de un mismo transcrito. También es el resultado de algunas mutaciones puntuales, aquellas en las que se insertan o eliminan nucleótidos en número no múltiplo de 3 (pág. 350).

22.1.4 Papel adaptador del tRNA

Debe recordarse que la "lectura" de la secuencia de codones está protagonizada por los tRNAs (págs. 71-75), que ejercen el papel de **adaptadores** de la información genética a la secuencia de las proteínas. Intervienen para ello como estructuras activas especializadas (formando los aminoacil-tRNAs), que dirigen los aminoácidos hacia el mRNA situado en el ribosoma. En este orgánulo tiene lugar la interacción del triplete de bases del aminoacil-tRNA (**anticodón**) con el triplete de bases del mRNA (**codón**) determinando, durante la traducción, la incorporación del aminoácido específico frente a cada posición del mRNA. Por tanto, coloquialmente, como ocurre de aquí en adelante, se dice simplemente que "un codón determina o codifica un aminoácido".

El tRNA es el mediador entre el mensaje del mRNA y la secuencia de aminoácidos del polipéptido que se está sintetizando:



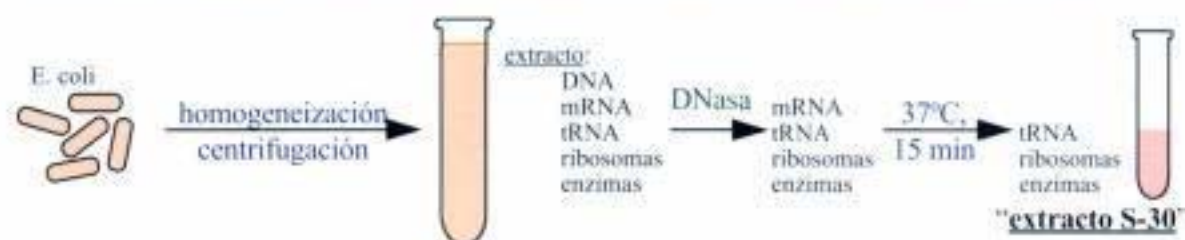
22.2 ASIGNACIÓN DE CODONES A AMINOÁCIDOS CONCRETOS

22.2.1 Traducción de polinucleótidos sintéticos

22.2.1.1 Ensayo con homopolímeros

La necesidad de un triplete para codificar un aminoácido se demostró inicialmente *in vitro* con el empleo de dos tipos de componentes:

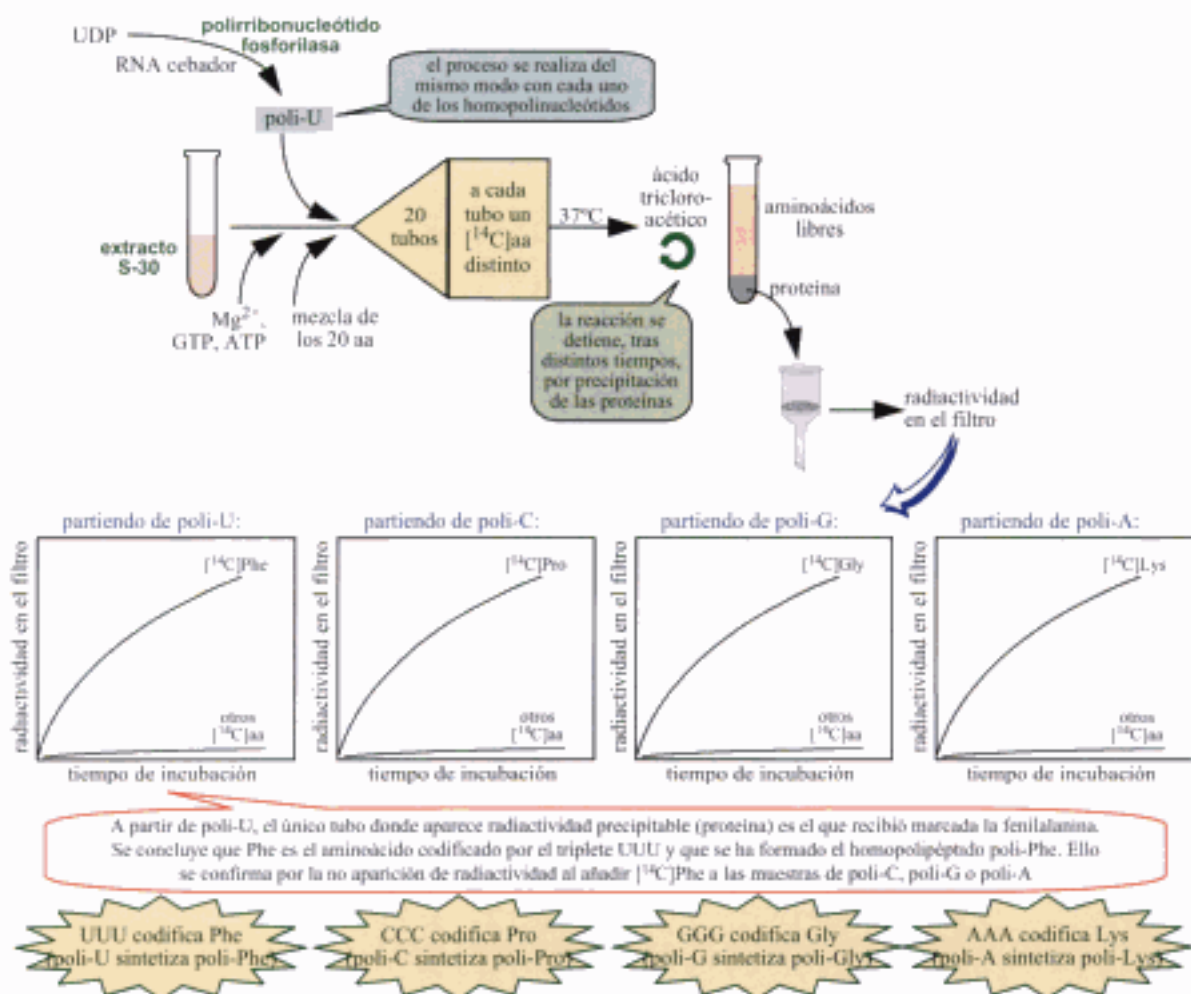
- Un *extracto libre de células* (inicialmente de *E. coli* o de reticulocitos de conejo, llamado "extracto S-30"), que aporta los componentes necesarios para que se desarrolle la síntesis proteica, a excepción del mRNA.



- Un homopolirribonucleótido sintético, por ejemplo poli-U, obtenido por catálisis con polinucleótido fosforilasa (pág. 246) a partir de UDP y un cebador, sin necesidad de molde.

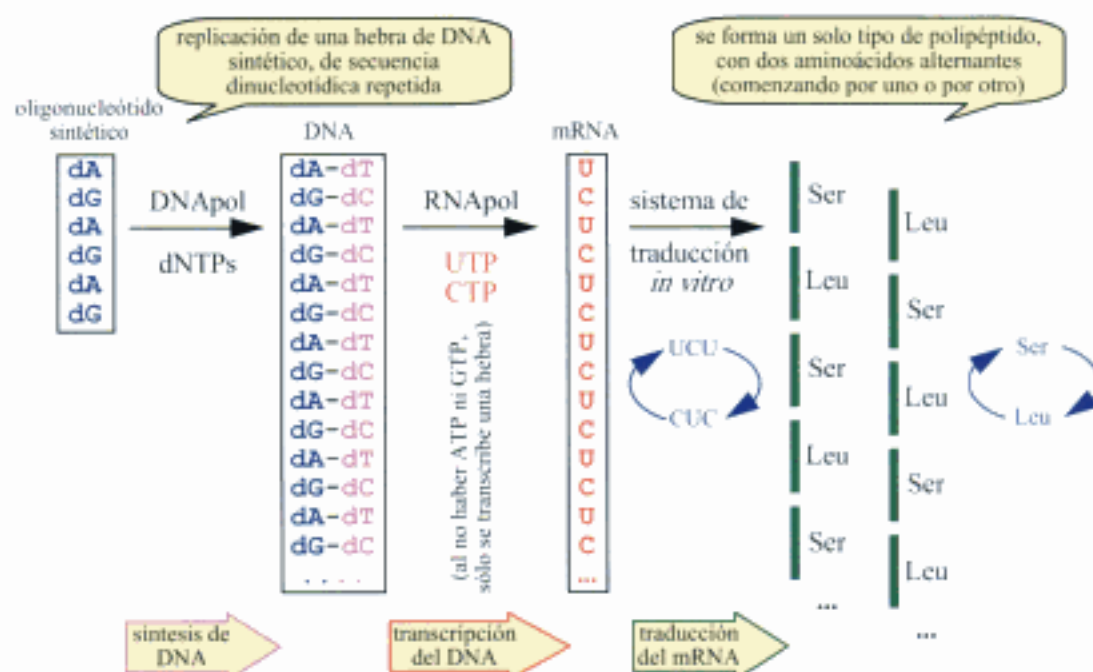


Al añadir este polirribonucleótido al extracto acelular (S-30) se constituye un sistema de traducción *in vitro*, produciendo un homopolipéptido. Empleando aminoácidos marcados con isótopos radiactivos, se puede averiguar cuál es el polipéptido obtenido.



Los experimentos anteriores sólo permiten descifrar algunos codones muy particulares; para desvelar la correspondencia entre los demás codones y los aminoácidos se emplearon polirribonucleótidos con secuencias repetidas de dos, tres o cuatro bases, que conducen a la síntesis de polipéptidos con uno o unos pocos aminoácidos en patrones repetitivos.

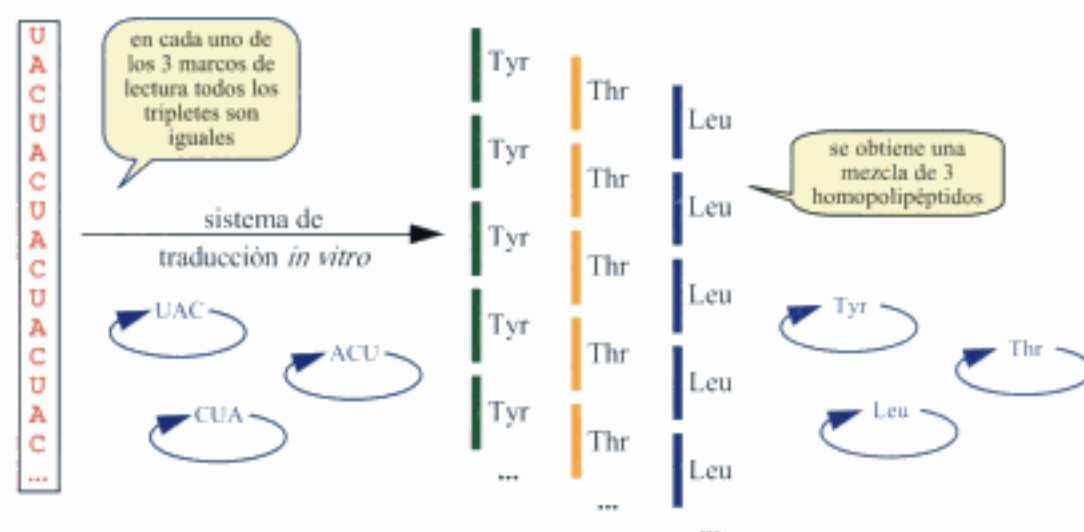
22.2.1.2 Ensayo con secuencias repetidas de dos nucleótidos



Se concluye que los tripletes UCU y CUC codifican Leu y Ser, o viceversa; con este ensayo no se puede definir la correspondencia de cada uno. Fueron experimentos posteriores los que demostraron que UCU codifica Ser y CUC codifica Leu.

22.2.1.3 Ensayo con secuencias repetidas de tres nucleótidos

Para simplificar, se representa directamente el mRNA, en este caso con una secuencia trinucleotídica repetida.



Se establece, pues, la correspondencia entre Tyr, Thr y Leu con UAC, ACU y CUA, de nuevo sin poder definir cuál de los codones codifica cada uno de los 3 aminoácidos.

22.3 MODELOS DE REPRESENTACIÓN

Se han elaborado diversos modelos (tablas, esquemas, etc.) para representar de forma clara y sencilla la correlación existente entre codones del mRNA y aminoácidos del polipéptido traducido. En todos ellos, la secuencia del triplete se refiere al codón del mRNA (de ahí la presencia de U y no de T), nunca se indica como triplete del DNA ni como triplete anticodón del tRNA (aunque se pueden y deben hacer referencias cruzadas entre las secuencias correspondientes). Las bases de cada codón se numeran, como siempre, en la dirección 5'→3', la misma en la que se sintetiza el mRNA; así, la 1ª posición es el extremo 5' del codón y la 3ª posición es su extremo 3'.

22.3.1 Formato en tabla

La representación más comúnmente usada utiliza filas y columnas, a modo de tabla de doble entrada. Pueden señalarse dos variantes:

		BASE INTERMEDIA (2ª BASE DEL CODÓN)							
		U		C		A		G	
BASE EN 5' (1ª BASE DEL CODÓN)	U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
		UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
		UUA	Leu	UCA	Ser	UAA	PARO	UGA	PARO
		UUG	Leu	UCG	Ser	UAG	PARO	UGG	Trp
	C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
		CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
		CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
		CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
	A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
		AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
		AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
		AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
	G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
		GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
		GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
		GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

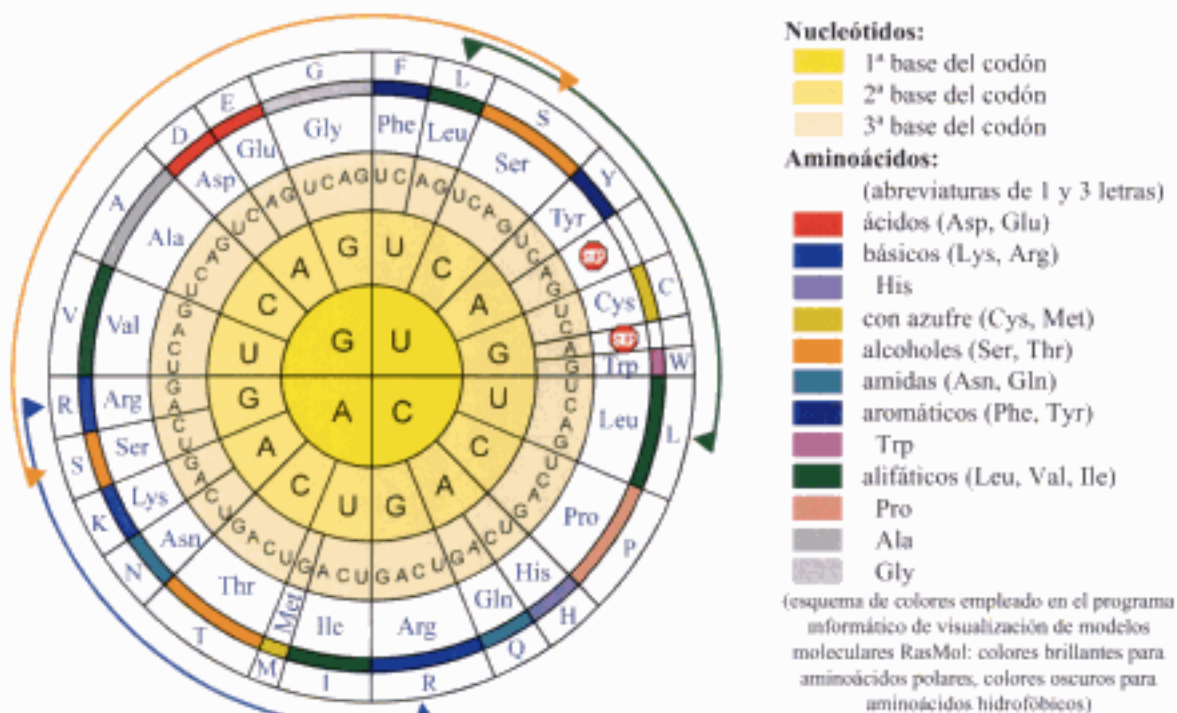
En la versión tradicional, las filas indican el 1º nucleótido y las columnas el 2º. Se obtienen 16 cuadrículas. Al añadir el 3º nucleótido aparecen 4 tripletes en cada cuadrícula, idénticos en sus dos primeros nucleótidos. Finalmente, se indica el aminoácido codificado por cada codón.

		BASE INTERMEDIA (2ª BASE DEL CODÓN)							
		U		C		A		G	
BASE EN 5' (1ª BASE DEL CODÓN)	U	Phe		Ser		Tyr		Cys	
		Leu				PARO		PARO	
								Trp	
	C			Pro		His		Arg	
						Gln			
	A	Ile		Thr		Asn		Ser	
		Met				Lys		Arg	
	G					Asp			
		Val		Ala				Gly	
						Glu			

En una versión algo más sencilla se evita la escritura completa del codón, el 1º nucleótido se escribe a la izquierda, el 2º arriba y el 3º a la derecha. En cada cuadrícula, sólo se escriben los aminoácidos codificados.

22.3.2 Formato circular

Un modelo más reciente, posiblemente más sencillo y fácil de utilizar, emplea círculos concéntricos en los que se indican, del centro hacia fuera, los 3 nucleótidos del codón y el aminoácido que codifica.



Representación construida con base en el "Diagrama del sol naciente para el código genético", portada de Biochemical Education, vol.23, nº3, julio 1995, Pergamon Press

22.4 CARACTERÍSTICAS ESPECÍFICAS

Se pueden analizar empleando las representaciones anteriores. Son las siguientes:

22.4.1 Codón de inicio

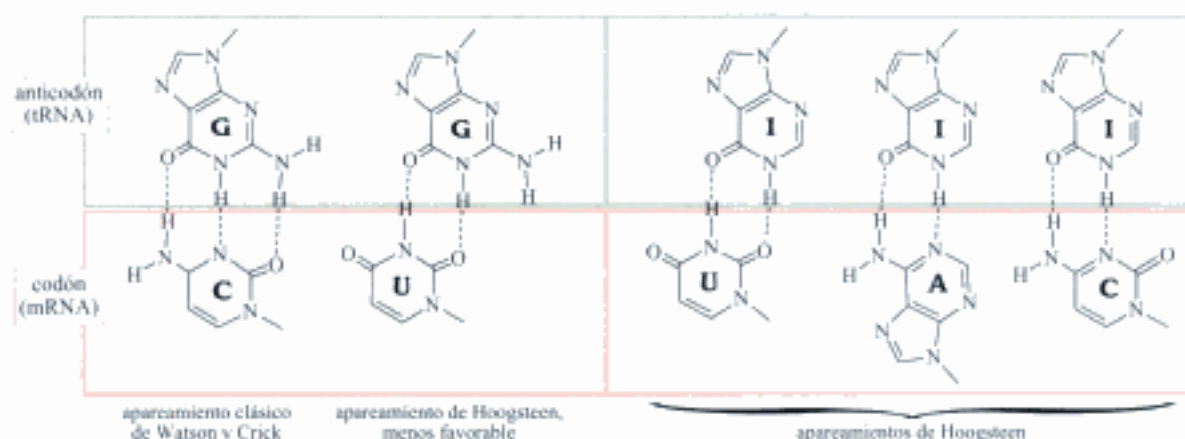
El codón **AUG**, el único que codifica metionina, es responsable como cualquier otro codón de la incorporación del aminoácido correspondiente al polipéptido, pero al mismo tiempo posee la particularidad de actuar de forma específica como **codón de inicio** en la mayoría de los casos (con mucha menos frecuencia se usan como inicio otros codones, en especial ACG, CUG y GUG). Es decir, la síntesis de la mayoría de las proteínas comienza en un codón AUG; ello explica que Met siempre ocupe la posición 1 (inicial o N-terminal) de cualquier polipéptido recién sintetizado. Como se verá luego (pág. 308), las proteínas eucarióticas poseen una metionina N-terminal, mientras que en procariotas se incorpora como formil-metionina. Para que AUG actúe como codón de inicio en eucariotas se requiere que forme parte de una secuencia específica (secuencia de Kozak, pág. 308).

22.4.2 Codones de terminación

Se puede observar en las representaciones del código genético que los codones **UAA**, **UAG** y **UGA** no codifican aminoácido alguno, sino que causan la finalización de la síntesis proteica (pág. 314). Se les denomina **codones de terminación**, **de paro**, **codones "stop"**, o también **codones sin sentido**.

La presencia de 3 codones de terminación en un total de 64 indicaría que, estadísticamente, la síntesis no se alargaría más allá de unos 20 aminoácidos ($3/64 = 1/21 = 4,7\%$). Sin embargo, la región estructural de los genes para proteínas contiene una frecuencia mucho menor de codones de terminación. La región del DNA comprendida entre un codón de inicio y uno de terminación se llama "marco abierto de lectura" (*open reading frame*, ORF). Por ejemplo, para una proteína de masa 60 kDa se requeriría un marco abierto de lectura con unos 545 codones (masa media de un residuo aminoácido = 110 Da; $60.000 \text{ Da} / 110 \text{ Da} = 545$).

Los detalles moleculares de las interacciones de Watson y Crick (pág. 41) y las menos favorables (enlaces de Hoogsteen, págs. 60 y 73) son los siguientes:



Resumiendo, los apareamientos posibles mediante "balanceo" son los siguientes:

1 ^{er} nucleótido (5') en el anticodón (tRNA):	3 ^{er} nucleótido (3') en el codón (mRNA):	resultado:
C	G (Watson-Crick)	C reconoce un solo codón
A	U (Watson-Crick)	A reconoce un solo codón
U	A (Watson-Crick)	U reconoce dos codones
	G (más débil)	
G	C (Watson-Crick)	G reconoce dos codones
	U (más débil)	
I	A (más débil)	I reconoce tres codones
	U (más débil)	
	C (más débil)	

22.4.3.2 Consideraciones prácticas

Todo lo anterior determina que el anticodón de un tRNA puede leer uno, dos o tres codones de un mRNA, en los cuales las 2 primeras bases son idénticas. Veamos ahora cómo puede calcularse el número de anticodones, el número de moléculas diferentes de tRNA necesarias para los 61 codones codificadores de los 20 aminoácidos. Aunque en las células ésta es una cifra imprecisa, desde un planteamiento teórico, basado en el aprovechamiento al máximo del balanceo, puede afirmarse que se necesitan *al menos* 31 tRNAs para "leer" todo el código.

codones
5'→3'

anticodones
3'→5'

tRNAs

Los cuatro codones de una cuadrícula precisan siempre dos anticodones, bien a costa de los balances de U y G o bien del balance de I

	U	C	A	G
U	UUU } AAG } tRNA ^{Phe}	UCU } AGG } tRNA ^{Ser}	UAU } AUG } tRNA ^{Tyr}	UGU } ACA } tRNA ^{Cys}
	UUC } AAU } tRNA ^{Leu}	UCC } AGU } tRNA ^{Ser}	UAC } UAA } PARO	UGC } UGA } PARO
C	CUU } GAG } tRNA ^{Leu}	CCU } GGG } tRNA ^{Pro}	CAU } GUG } tRNA ^{His}	CGU } GCG } tRNA ^{Arg}
	CUC } GAU } tRNA ^{Leu}	CCC } GGC } tRNA ^{Pro}	CAC } GUU } tRNA ^{Glu}	CGC } GCU } tRNA ^{Arg}
A	CUA } GAU } tRNA ^{Leu}	CCA } GGU } tRNA ^{Pro}	CAA } GUA } tRNA ^{Glu}	CGA } GCU } tRNA ^{Arg}
	CUG } GAC } tRNA ^{Leu}	CCG } GGC } tRNA ^{Pro}	CAG } AUU } tRNA ^{Asn}	CGG } GCU } tRNA ^{Arg}
G	AUU } UAU } tRNA ^{Ile}	ACU } UGG } tRNA ^{Thr}	AAU } AUG } tRNA ^{Asn}	AGU } UUG } tRNA ^{Ser}
	AUC } UAU } tRNA ^{Ile}	ACC } UGU } tRNA ^{Thr}	AAC } AAU } tRNA ^{Lys}	AGC } UCC } tRNA ^{Arg}
G	AUA } UAU } tRNA ^{Ile}	ACA } UGU } tRNA ^{Thr}	AAA } UUU } tRNA ^{Lys}	AGA } UCC } tRNA ^{Arg}
	AUG } UAU } tRNA ^{Met}	ACG } UGU } tRNA ^{Thr}	AAG } GAU } tRNA ^{Asp}	AGG } GGU } tRNA ^{Gly}
G	GUU } CAG } tRNA ^{Val}	GCU } CCG } tRNA ^{Ala}	GAU } CUG } tRNA ^{Asp}	GGU } CCG } tRNA ^{Gly}
	GUC } CAU } tRNA ^{Val}	GCC } CCG } tRNA ^{Ala}	GAC } GAG } tRNA ^{Glu}	GGC } GUC } tRNA ^{Gly}
G	GUA } CAU } tRNA ^{Val}	GCA } CGU } tRNA ^{Ala}	GAA } CUU } tRNA ^{Glu}	GGA } CCU } tRNA ^{Gly}
	GUG } CAC } tRNA ^{Val}	GCG } CGU } tRNA ^{Ala}	GAG } CUU } tRNA ^{Glu}	GGG } CCU } tRNA ^{Gly}

Los anticodones con G como 1ª base (5') aparecen por balanceo con 2 codones, los terminados en U y C.
 Los anticodones con U en 1ª posición (5') aparecen por balanceo con los 2 codones cuya 3ª base (3') es A y G.

Los anticodones con inosina (I) como 1ª base aparecen por balanceo con 3 codones, los terminados en U, C y A.
 Se requiere otro anticodón para el 4º codón de la cuadrícula, sin balanceo.

nº de codones por aminoácido	nº de anticodones por aminoácido	nº total de aminoácidos	nº total de anticodones	nº total de codones
1	1	2	2	2
2	1	9	9	18
3	1	1	1	3
4	2	5	10	20
6	3	3	9	18
suma:		20	31	61

22.4.4 Universalidad

Los primeros estudios indicaron que la correspondencia entre codones y aminoácidos era la misma independientemente del organismo estudiado, conduciendo a la afirmación de que "el código genético es **universal**". Actualmente se sabe que esto no es completamente cierto; se han encontrado excepciones, aunque escasas, en las mitocondrias humanas y de otros mamíferos y en ciertas bacterias. Debemos, pues, decir que **el código genético es casi universal**.

	núcleo	mitocondrias
AGA, AGG	Arg	paro
AUA	Ile	Met
UGA	paro	Trp

Por tanto, comparten el mismo código genético los procariotas, los núcleos de todos los eucariotas y las mitocondrias de plantas. Parece probable que todas las formas de vida hayan tenido un antepasado evolutivo común, con un solo código genético, que se ha conservado muy bien a todo lo largo de la evolución biológica. Las diferencias pueden haber surgido durante la evolución por transferencia del genoma procariótico original, con código genético universal, al genoma mitocondrial.

Tema 23

Síntesis de proteínas: Traducción

23.1 Características de la traducción	300
23.1.1 Sentido de avance de la síntesis proteica	300
23.1.2 Carácter monocistrónico o policistrónico del mRNA	301
23.1.3 Fases de la traducción	302
23.2 Fase previa: activación de los aminoácidos en forma de aminoacil-tRNAs	302
23.2.1 Reacciones de aminoacilación	302
23.2.2 Aminoacil-tRNA sintetasas	304
23.2.2.1 Tipos de sintetasas	304
23.2.2.2 Especificidad de las aminoacil-tRNA sintetasas	304
23.2.2.3 Mecanismo de reconocimiento de los aminoácidos por las sintetasas	304
23.2.2.4 Mecanismo de reconocimiento de los tRNAs por las sintetasas	304
23.2.3 Corrección de errores cometidos por las sintetasas	305
23.2.4 Balance energético de la activación	307
23.3 Fase 1: iniciación	307
23.3.1 Especificidad del punto de inicio	307
23.3.2 Formación del metionil-tRNA iniciador	308
23.3.3 Pasos sucesivos en la fase de inicio	309
23.3.3.1 Disociación del ribosoma (paso 1)	309
23.3.3.2 Unión del aminoacil-tRNA y del mRNA a la subunidad menor del ribosoma: complejo de preiniciación (paso 2)	310
23.3.3.3 Unión de la subunidad mayor: complejo de iniciación (paso 3)	310
23.3.3.4 Reutilización de los factores de iniciación	310
23.3.3.5 Balance energético de la iniciación	310
23.4 Fase 2: elongación o alargamiento de la cadena peptídica	311
23.4.1 Ubicación del aminoacil-tRNA en el sitio A (paso 4)	311
23.4.2 Transpeptidación: formación del enlace peptídico (paso 5)	313
23.4.3 Translocación: desplazamiento del ribosoma en un codón (paso 6)	313
23.4.4 Repetición del ciclo de elongación (pasos 4, 5 y 6)	314
23.5 Fase 3: terminación	314
23.5.1 Unión del factor de liberación (paso 7)	314
23.5.2 Hidrólisis del peptidil-tRNA (paso 8)	315
23.5.3 Disociación (paso 9)	315
23.6 Energética de la síntesis de proteínas	315
23.7 Inhibidores de la traducción	316
23.8 Regulación de la síntesis proteica: Niveles de control	318
23.8.1 Control del transporte de mRNA	318
23.8.2 Control de la capacidad y frecuencia de traducción de los mRNAs	319
23.8.2.1 Ejemplo 1: Control de la síntesis de ferritina por bloqueo del mRNA	319
23.8.2.2 Ejemplo 2: Regulación de la síntesis de hemoglobina por fosforilación de un factor de iniciación	320

23.1 CARACTERÍSTICAS DE LA TRADUCCIÓN

La traducción consiste en la síntesis de las proteínas mediante la unión de aminoácidos según el orden establecido por la secuencia de nucleótidos del mRNA y el código genético. Se completa así el "dogma central" de transmisión de la información genética a partir del DNA genómico original (págs. 1 y 244).

Puesto que este proceso es esencialmente igual en todo tipo de organismos (salvo pequeñas diferencias en los mecanismos y componentes implicados), nuestro estudio se centra en eucariotas, aludiendo a procariotas cuando lo requieran sus aspectos particulares en favor de una mejor comprensión del proceso en términos moleculares. Se plantea en estrecha relación con el código genético (tema 22) y se completará con las modificaciones postraduccionales (tema 24), necesarias para el ejercicio de su función; finalmente, se analizarán los mecanismos de degradación, de forma análoga a lo ya tradicional para otras biomoléculas (glúcidos y lípidos, principalmente). Con todo ello se cierran los dos grandes niveles, biosíntesis y degradación, del metabolismo proteico. Este enfoque es imprescindible hoy día para conocer el funcionamiento normal y patológico de las proteínas (base de las enfermedades moleculares, tema 28) y para abordar en el futuro los avances en la naciente Proteómica (pág. 217) con vistas al diagnóstico y la terapéutica.

Desde un punto de vista bioquímico, la traducción se caracteriza por la gran variedad de proteínas formadas, su elevado coste energético (consume un 80-90% de la energía de biosíntesis) y la necesidad de una regulación muy estrecha en respuesta a las necesidades celulares y al ritmo de degradación proteica. Es, posiblemente, el más complejo de los procesos de síntesis, en el que participa un mayor número de macromoléculas diferentes. Las principales son:

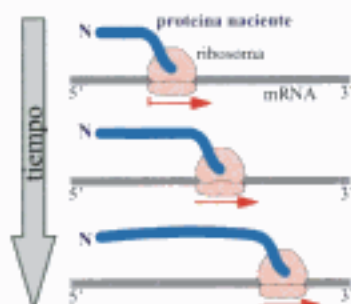
- Un tRNA especial para la iniciación (tRNA_{Met} en procariotas, tRNA_{Met} en eucariotas; pág. 308).
- 31 tipos de tRNAs portadores de aminoácidos, en forma de aminoacil-tRNAs. En realidad, éste es el número mínimo necesario teóricamente (págs. 298 y 304-305), en las células suele haber un número superior de tRNAs diferentes; por ejemplo, 40 es una cifra frecuente y en la cepa K12 de *E. coli* se han encontrado 84 tRNAs.
- Ribosomas, formados por numerosos rRNAs y proteínas (pág. 80).
- Un mRNA, de secuencia distinta para cada polipéptido sintetizado.
- Factores proteicos de inicio, elongación y terminación de la síntesis.

En una célula existen millares de estos participantes: unos 20.000 ribosomas, 100.000 moléculas de enzimas y factores proteicos, 200.000 de tRNAs, etc. Se requiere, además, un acervo (*pool*) de al menos 20 aminoácidos, ATP, enzimas aminoacil-sintetasas para activar los aminoácidos, GTP para varias etapas, grupos reductores, iones (Mg^{+2} y K^{+}), enzimas auxiliares, etc. Intervienen, directa o indirectamente, casi todas las estructuras subcelulares (núcleo, citoplasma, orgánulos y membranas). A pesar de ello, el proceso es muy rápido: en bacterias, la velocidad de traducción es de unos 20 aminoácidos por segundo, mientras que en eucariotas es más reducida, se incorporan de 2 a 4 aminoácidos por segundo.

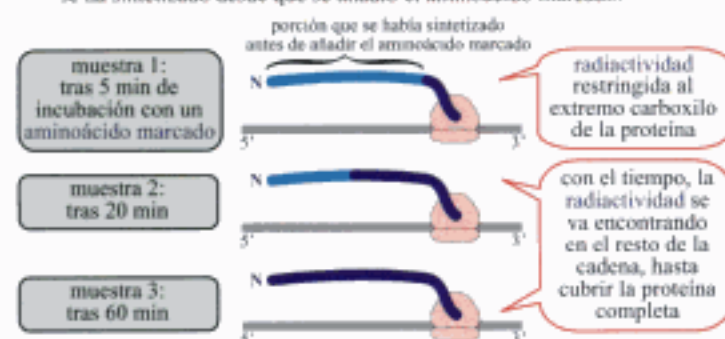
23.1.1 Sentido de avance de la síntesis proteica

El primer aspecto básico, el sentido en el que se realiza la síntesis, puede ser considerado desde un doble punto de vista: la secuencia nucleotídica del mRNA se lee de 5' a 3', mientras que la secuencia aminoacídica del polipéptido se incorpora desde el extremo amino hacia el carboxilo. Esto último se demostró por primera vez mediante incubación de reticulocitos (precursores de eritrocitos que sintetizan hemoglobina) con aminoácidos marcados, midiendo la radiactividad en la globina aislada tras distintos tiempos de marcaje.

- 1 El proceso de crecimiento del polipéptido avanza así:



- 2 Al estudiar la proteína completa a distintos tiempos tras haber añadido un aminoácido marcado, sólo presenta marcaje la parte de proteína que se ha sintetizado desde que se añadió el aminoácido marcado:



- 3 Localización del marcaje en la proteína completa a distintos tiempos:



- 4 **Conclusión:** la síntesis del polipéptido comienza por su extremo N-terminal, y los aminoácidos se van añadiendo al extremo C-terminal del primer aminoácido y del polipéptido naciente. Es decir, cada aminoácido se incorpora uniéndose su grupo amino al carboxilo del aminoácido anterior.

23.1.2 Carácter monocistrónico o policistrónico del mRNA

Un aspecto importante en relación con la traducción es que un mRNA puede dar lugar a la síntesis de un polipéptido o de varios. Esta propiedad se define mediante el concepto de **cistrón**, que conviene distinguir del de **gen**, aunque están estrechamente relacionados y ambos pueden considerarse una "unidad" de funcionalidad genética.

- Los **mRNAs monocistrónicos**, presentes en procariotas y típicos de eucariotas, codifican *un solo polipéptido*. Ello implica que la traducción se inicia en un solo sitio del mRNA, el codón de inicio, para finalizar en otro sitio, un codón de terminación.
- Los **mRNAs policistrónicos**, que sólo aparecen en procariotas y virus, codifican *varios polipéptidos*, por lo que la traducción se inicia (simultáneamente o no) en varios sitios, en varios codones de inicio, para finalizar en sendos codones de terminación.

mRNA procariótico

Caso nº 1: **monocistrónico**; un sitio de inicio de la traducción



Caso nº 2: **policistrónico**; varios sitios de inicio de la traducción (lo más frecuente)



mRNA eucariótico

Siempre **monocistrónico**; un sitio de inicio de la traducción

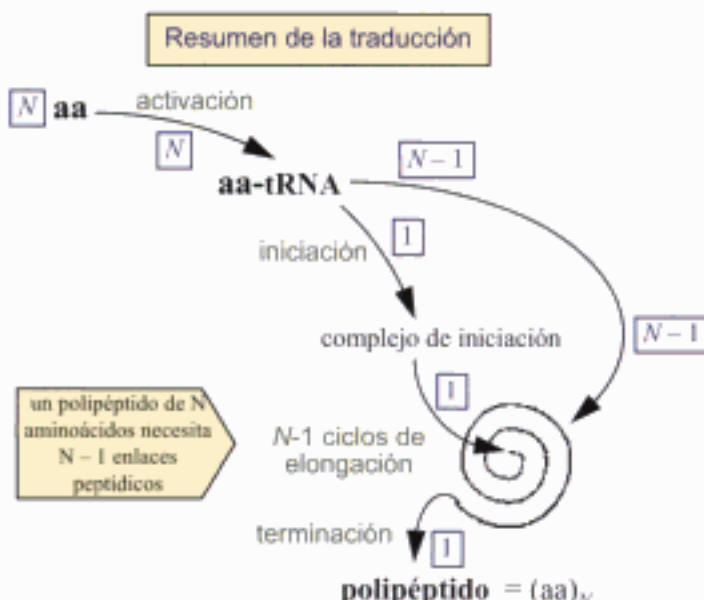


El concepto de **gen** (pág. 244) corresponde a una región codificante del DNA genómico, responsable de un producto génico, sea éste RNA (por transcripción y maduración) o proteína (por transcripción, maduración, traducción y modificación postraduccional). En el caso de mRNAs monocistrónicos, el concepto de **cistrón** es prácticamente indistinguible del de gen. La diferencia importante surge en los genes que se transcriben dando un mRNA policistrónico (y, por tanto, sólo afecta a los organismos procariotas).

El concepto de **cistrón**, en su forma genérica, se diferencia: a) por ser más amplio que el de gen, al considerar como región codificante tanto al DNA como al RNA (ambos dan lugar a un producto génico), y b) por ser al mismo tiempo más estricto, pues se define como la región del genoma (DNA) o del mRNA que codifica un producto génico único. Normalmente se emplea sólo para este último caso, por lo que un **cistrón** es la secuencia de mRNA que codifica una molécula de polipéptido. Recuérdese que la definición más amplia de **gen** obligaba a incluir casos límite como consecuencia de la maduración postraduccional (pág. 276); de forma similar ocurre en el caso que nos ocupa, el mRNA policistrónico, con una unidad de transcripción que codifica varios productos (conocida también habitualmente como **operón**). Por lo tanto, en este caso se trata de un *gen policistrónico*.

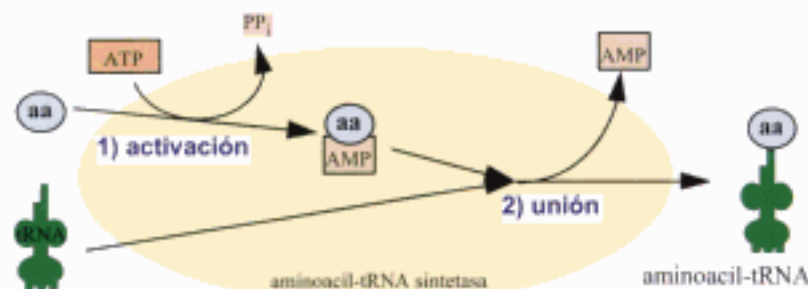
23.1.3 Fases de la traducción

Dentro del proceso de traducción se pueden distinguir varias etapas o fases: una **previa o de activación** de los aminoácidos precursores y tres bien diferenciadas, **iniciación**, **elongación** y **terminación**, bajo un criterio análogo al seguido para la replicación y la transcripción (temas 12 y 19). La maduración postraduccional, considerada a veces como 5ª fase, se trata hoy día de modo independiente (tema 24).



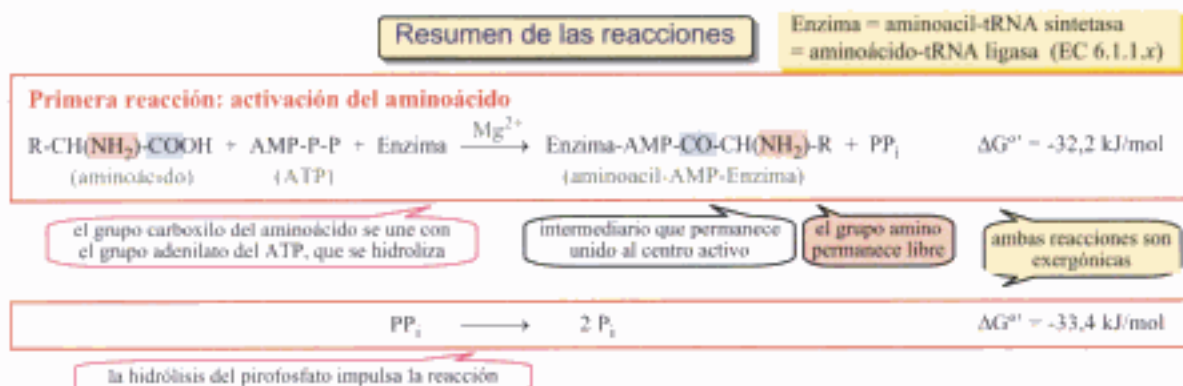
23.2 FASE PREVIA: ACTIVACIÓN DE LOS AMINOÁCIDOS EN FORMA DE AMINOACIL-tRNAs

La unión a cada tRNA del aminoácido correcto (el que, según el código genético, está codificado por el codón complementario del anticodón que posee ese tRNA) es posiblemente el paso más crítico de la expresión génica. Si no se hace de forma correcta, carece de sentido todo el código genético. Esa unión ocurre en el citosol celular, como el resto del proceso de traducción, y supone la activación de los aminoácidos (en forma de aminoacil-tRNAs) para poder formar enlaces peptídicos. De este forma, la fase de activación define el papel conector del tRNA con los aminoácidos en la traducción.



23.2.1 Reacciones de aminoacilación

La unión del aminoácido con el tRNA se verifica mediante dos reacciones, ambas catalizadas por una misma enzima (en un centro activo único o en dos centros próximos). Esta enzima es una de las **aminoacil-tRNA sintetetas** (abreviadamente, sintetetas o **aaRSs**).



Segunda reacción: unión con el tRNA

el grupo aminoacilo se transfiere desde el intermediario hacia el tRNA

el grupo amino permanece libre

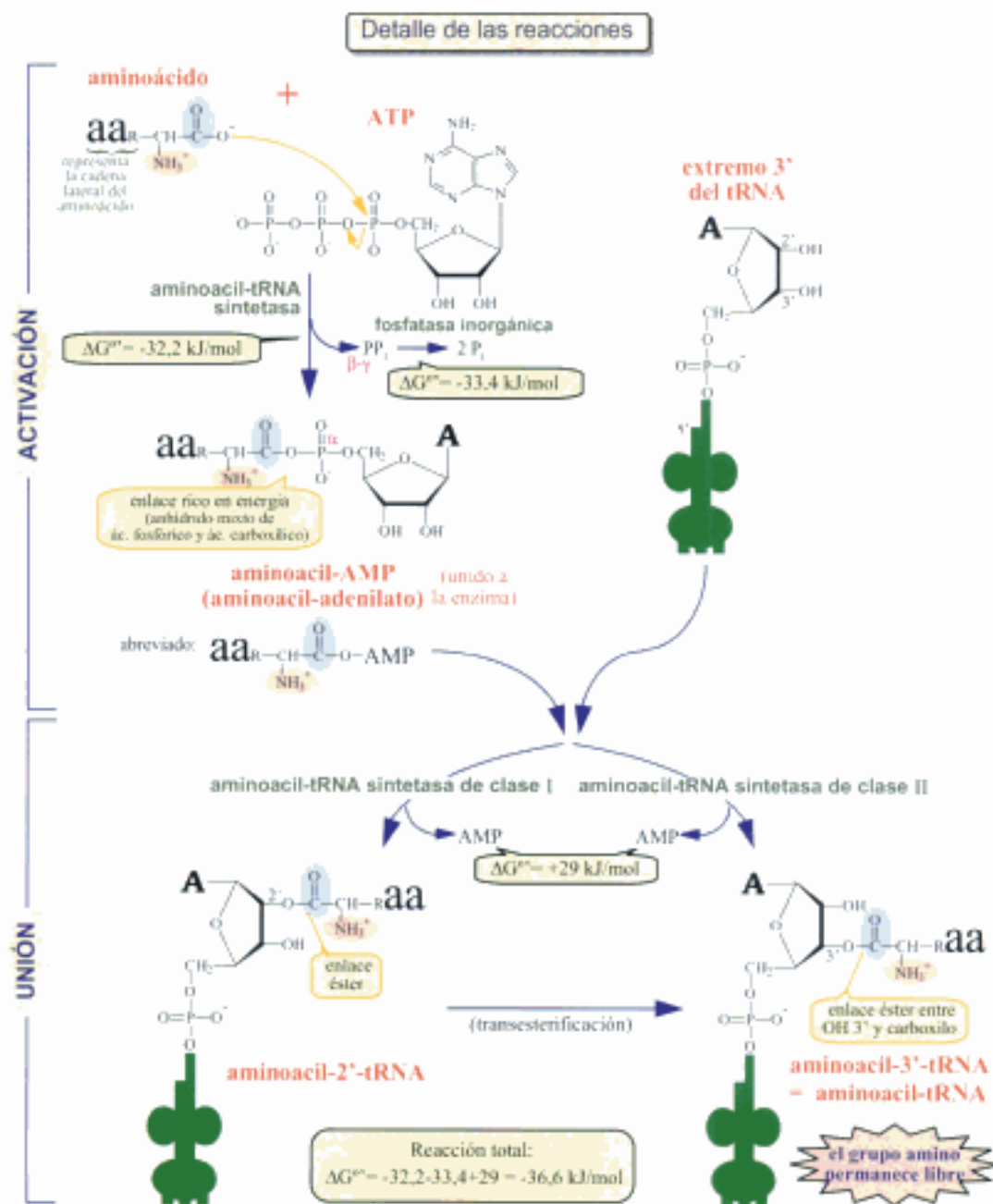
la reacción es endérgica

Reacción total:

el aminoacil-tRNA es una forma activada, de alta energía, del aminoácido

la reacción neta es exérgica

El producto final de la doble reacción es un **aminoacil-tRNA**, con un enlace éster entre el grupo OH 3' terminal del tRNA y el grupo carboxilo del aminoácido (véase pág. 71 para la forma de nombrarlo). En el caso de las aminoacil-tRNA sintetasas de clase I (pág. 304), el éster se forma inicialmente sobre el grupo OH 2', pero luego se convierte al isómero 3'.



23.2.2 Aminoacil-tRNA sintetetasas

23.2.2.1 Tipos de sintetetasas

Se han aislado multitud de aminoacil-tRNA sintetetasas de bacterias, plantas y animales, cuya masa oscila entre 88 y 181 kDa. En algunos casos se han podido secuenciar y cristalizar para estudiar su estructura. Todas las sintetetasas catalizan de forma similar la reacción de activación del aminoácido con AMP, pero presentan diferencias en la segunda reacción, la de unión al tRNA. Concretamente, unas unen el aminoácido al grupo hidroxilo 2' y otras al 3', siempre del último nucleótido del tRNA (pág. 303). De acuerdo con esto, y también con diferencias estructurales, se clasifican las sintetetasas en dos grupos denominados, respectivamente, **clase I** y **clase II**.

tipo de aaRS	grupo OH del tRNA	transesterificación posterior	estructura de la enzima	aminoácidos que reconocen
de clase I	2'	es necesaria	la mayoría monoméricas	Arg Cys Gln Glu Ile Leu Met Trp Tyr Val (10, los más grandes y polares)
de clase II	3'	no se requiere	la mayoría diméricas	Ala Asn Asp Gly His Lys Phe Pro Ser Thr (10, los más pequeños e hidrófobos)

23.2.2.2 Especificidad de las aminoacil-tRNA sintetetasas

Ya se ha explicado (pág. 296) cómo la hipótesis del balanceo permite que un mínimo de 31 tRNAs reconozcan los 61 codones que codifican los 20 aminoácidos. Se forman, pues, 32 aminoacil-tRNAs diferentes (debe añadirse un tRNA específico para la incorporación de la metionina de iniciación). Cada uno de éstos interaccionará, a través de su anticodón, con el codón del mRNA que codifica el aminoácido correspondiente. En muchos casos, un mismo tRNA puede interaccionar con varios codones sinónimos, gracias al balanceo. Por otra parte, el grupo de codones (4 ó 6) de algunos aminoácidos requiere más de un tRNA; éstos se llaman **tRNAs sinónimos** o **isoaceptores** (pág. 71).

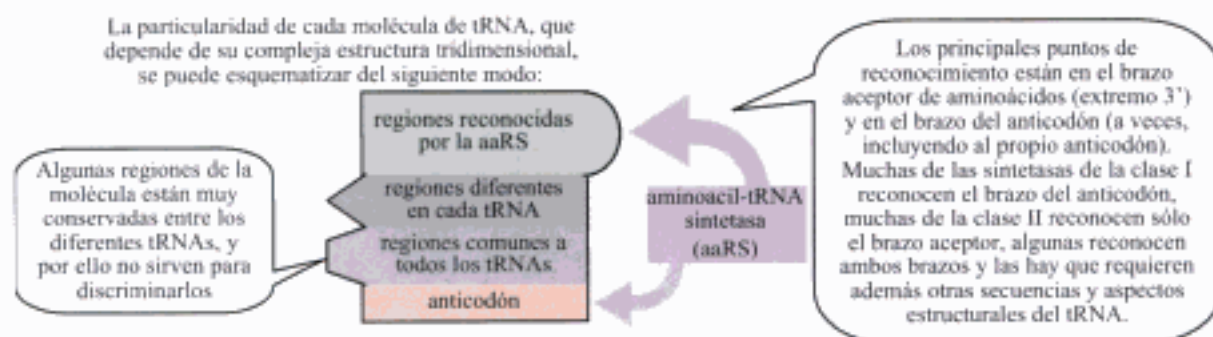
Este papel adaptador de los tRNAs entre la información genética del mRNA (nucleótidos) y la secuencia del polipéptido (aminoácidos) sólo puede desempeñarse gracias a la especificidad de las sintetetasas por ambos sustratos, el tRNA y el aminoácido. Cada aminoacil-tRNA sintetasa reconoce un único aminoácido (del que toma el nombre) y también todos los tRNAs cuyos anticodones codifican ese mismo aminoácido (codones sinónimos en el código genético). De acuerdo con ello, existen en la célula 20 aaRSs distintas. El papel de las sintetetasas es, por tanto, esencial para la fidelidad de la síntesis proteica y constituye la esencia del código genético, el problema central de la traducción; por ello, se suele considerar que las aaRSs representan un **segundo código genético**.

23.2.2.3 Mecanismo de reconocimiento de los aminoácidos por las sintetetasas

La interacción específica de un solo aminoácido con cada sintetasa es análoga al reconocimiento de cualquier otro sustrato por la enzima correspondiente. Cada aminoácido encaja en un determinado lugar del centro activo de la sintetasa, gracias a interacciones por puentes de hidrógeno, electrostáticas e hidrofóbicas. Sólo se unirán aquellos aminoácidos con un número suficiente de interacciones favorables.

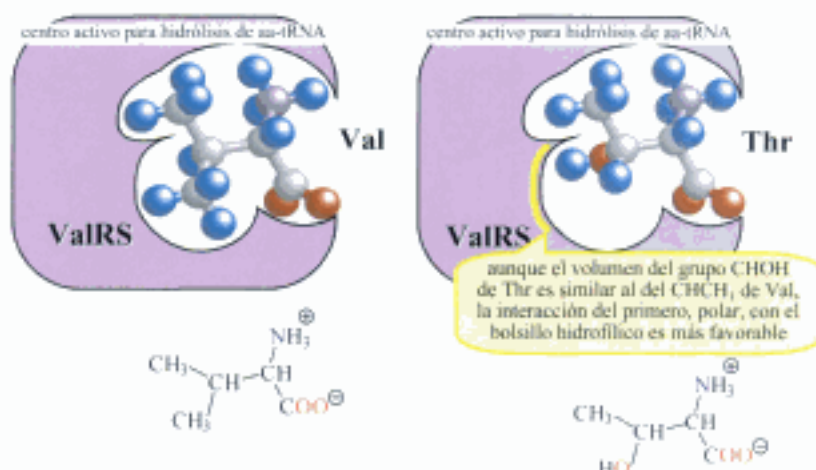
23.2.2.4 Mecanismo de reconocimiento de los tRNAs por las sintetetasas

La capacidad de una sintetasa, específica para un único aminoácido, para reconocer varios tRNAs, y al mismo tiempo rechazar otros, depende del modo como éstos interaccionan con el centro activo de la aaRS. Este reconocimiento es complejo y diferente para cada pareja tRNA-sintetasa, implicando no sólo las variaciones estructurales particulares de cada enzima (en especial, según que sean de clase I o II), sino también la interacción con diversas zonas de la molécula de tRNA.



Ejemplo 2. Con el mismo caso anterior, si se ha llegado a formar erróneamente Val-AMP-E en el sitio activo de IleRS, se percibe el error y se hidroliza el anhidrido entre Val y AMP. Es probable que esta corrección de pruebas tenga lugar porque el sitio hidrolítico de la enzima admite a Val-AMP (incorrecto) pero excluye o rechaza, por impedimento estérico, Ile-AMP (correcto, de mayor tamaño).

Ejemplo 3. La enzima valil-tRNA sintetasa (ValRS) sirve para demostrar la capacidad de corregir errores entre dos aminoácidos cuyo tamaño es casi idéntico por tener cadenas laterales muy similares; concretamente, Val y Thr. ValRS rechaza a Thr por una combinación de dos mecanismos. El centro activo de transferencia es hidrofóbico e interacciona más favorablemente con Val, de cadena lateral más hidrofóbica. Por el contrario, el centro activo de hidrólisis es hidrofílico e interacciona mejor con Thr, más polar. Como resultado, la Thr unida por error a tRNA^{Val} se hidrolizará preferentemente, mientras que Val unida correctamente a tRNA^{Val} se conserva.



A pesar de todos estos mecanismos de corrección, la tasa de error de la síntesis proteica es mucho mayor que la de, por ejemplo, la replicación (respectivamente, uno por cada 10⁴ aminoácidos y uno por cada 10⁹-10¹⁰ nucleótidos). Esto es permisible puesto que el error en una proteína desaparece al degradarse ésta y no pasa a futuras generaciones y supone un grado de fidelidad suficiente en la práctica, sin implicar un gasto energético excesivo.

23.1.1 Balance energético de la activación

El **primer gasto energético** que se debe considerar en la traducción es el necesario para la formación de aminoacil-tRNAs para **todos** los aminoácidos que componen la proteína. Como se ha descrito (pág. 302), en cada reacción de aminoacilación se hidrolizan dos enlaces fosfato de alta energía (ATP → AMP + PP_i y PP_i → 2 P_i) para formar el enlace éster entre aminoácido y tRNA.

No se tendrá en cuenta para el balance global, aunque evidentemente es significativo en la situación real en la célula, el gasto energético desperdiciado cada vez que un aminoácido se incorpora incorrectamente (pág. 306); cada error supone dos enlaces fosfato, al haberse sintetizado un aa-tRNA inútil, que es hidrolizado durante la corrección por la sintetasa.

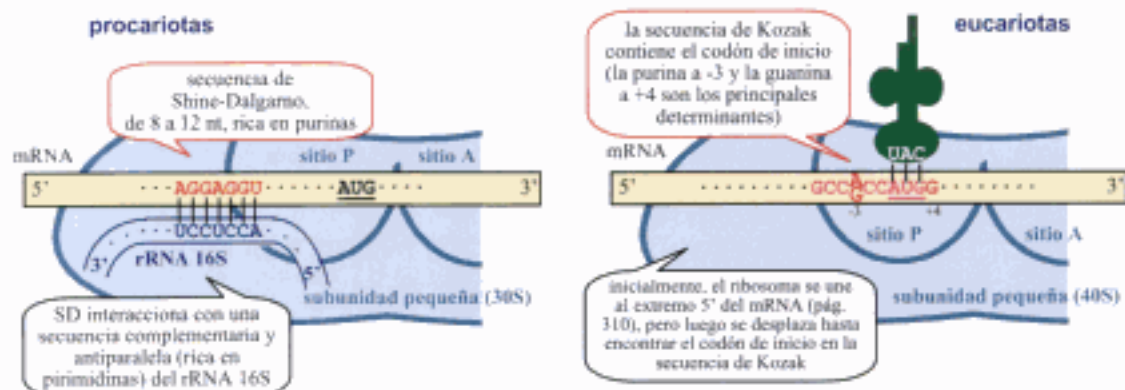
23.2 FASE 1: INICIACIÓN

Esta etapa, previa a la formación del primer enlace peptídico, determina normalmente la velocidad global de la síntesis. Transcurre de forma semejante en procariotas y eucariotas, salvo por la identidad de las proteínas implicadas, por lo que la explicación se centrará en estos últimos. Tanto en esta fase como en las de elongación y terminación es característica la intervención de complejos formados por las 3 moléculas principales –tRNAs, ribosomas (completos o una sola subunidad) y mRNAs–, además de otros factores proteicos, coenzimas, iones, etc. Para facilitar la descripción didáctica de estas fases, aunque sin existir consenso al respecto, se hace referencia a dichos complejos en función del número de moléculas principales que contienen: **complejos binarios** (dos moléculas) y **ternarios** (tres moléculas).

23.2.1 Especificidad del punto de inicio

La traducción nunca comienza por el extremo 5' del mRNA (por ello, se ha venido representando siempre una "región líder 5' no traducida", por ejemplo en pág. 301). Sólo un codón AUG (pág. 294), situado internamente en la molécula, actúa como punto de inicio, para lo cual necesita ser reconocido como tal por el ribosoma y el tRNA^{Met} iniciador.

En **procariotas**, cada cistron de un mRNA (pág. 301) posee una secuencia llamada **de Shine-Dalgarno (SD)**, que interacciona específicamente con la subunidad pequeña del ribosoma, definiendo la elección como codón de inicio de un AUG situado unos 10 nt corriente abajo.



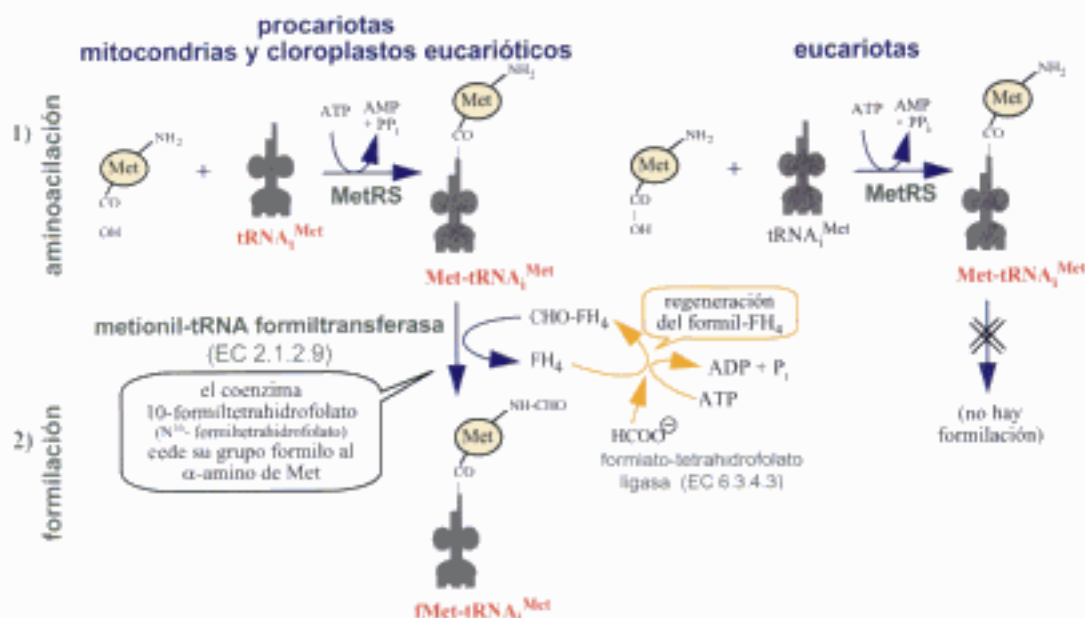
En **eucariotas**, el mRNA es casi siempre monocistónico (pág. 301), por lo que tiene un solo codón de inicio. No existe una señal especial que una el mRNA a la subunidad mayor del ribosoma, sino que la especificidad se establece a través de un tRNA especial, iniciador. La unión de éste al mRNA, en el entorno proporcionado por la subunidad menor del ribosoma, requiere que el codón AUG forme parte de la **secuencia de Kozak**, y da lugar a un complejo "ternario" mRNA : ribosoma : tRNA^{Met} (como se verá después, el tRNA participa realmente como metionil-tRNA^{Met} asociado al factor eIF-2 y a GTP).

23.3.2 Formación del metionil-tRNA iniciador

Los codones AUG en los que se inicia la traducción (residuo Met del extremo N-terminal del polipéptido) son reconocidos por un tRNA iniciador, llamado tRNA^{Met} (en procariotas se le suele llamar tRNA_f^{Met}, pero usaremos el primer nombre de forma genérica). Los codones AUG que corresponden a posiciones internas en la cadena polipeptídica interactúan con un tRNA distinto, que se llama simplemente tRNA^{Met}. En ambos casos el tRNA participa como metionil-tRNA^{Met} (etapa previa de activación del aminoácido, pág. 302).

La incorporación de metionina interna corresponde a la etapa de elongación, y tiene lugar de la misma forma que otros aminoácidos; la de metionina N-terminal ocurre, sin embargo, durante la iniciación, y de forma ligeramente diferente en procariotas y eucariotas:

En **procariotas** (y en **mitocondrias** y **cloroplastos** eucarióticos) se forma Met-tRNA^{Met}, que luego es modificado con un grupo formilo en el α-amino de la metionina. Es decir, hay **aminoacilación** y **formilación**, para incorporar N-formil-metionina (fMet) al inicio de la traducción. Como consecuencia, todas las proteínas recién sintetizadas comienzan por formilmetionina. Debe resaltarse que no se utiliza fMet como sustrato, ni tampoco un fMet-tRNA^{Met}. Esta etapa supone un pequeño coste energético adicional para la traducción (que no se contemplará al hacer el balance general, pág. 315): aunque la reacción no consume energía directamente, sí lo hace la regeneración del coenzima formiltetrahidrofolato.



- En la traducción citoplasmática de **eucariotas** no hay participación de formilmetionina, sólo se requiere la **aminoacilación** para incorporar metionina al inicio de la síntesis. Por tanto, los polipéptidos formados contienen metionina en su extremo N-terminal.

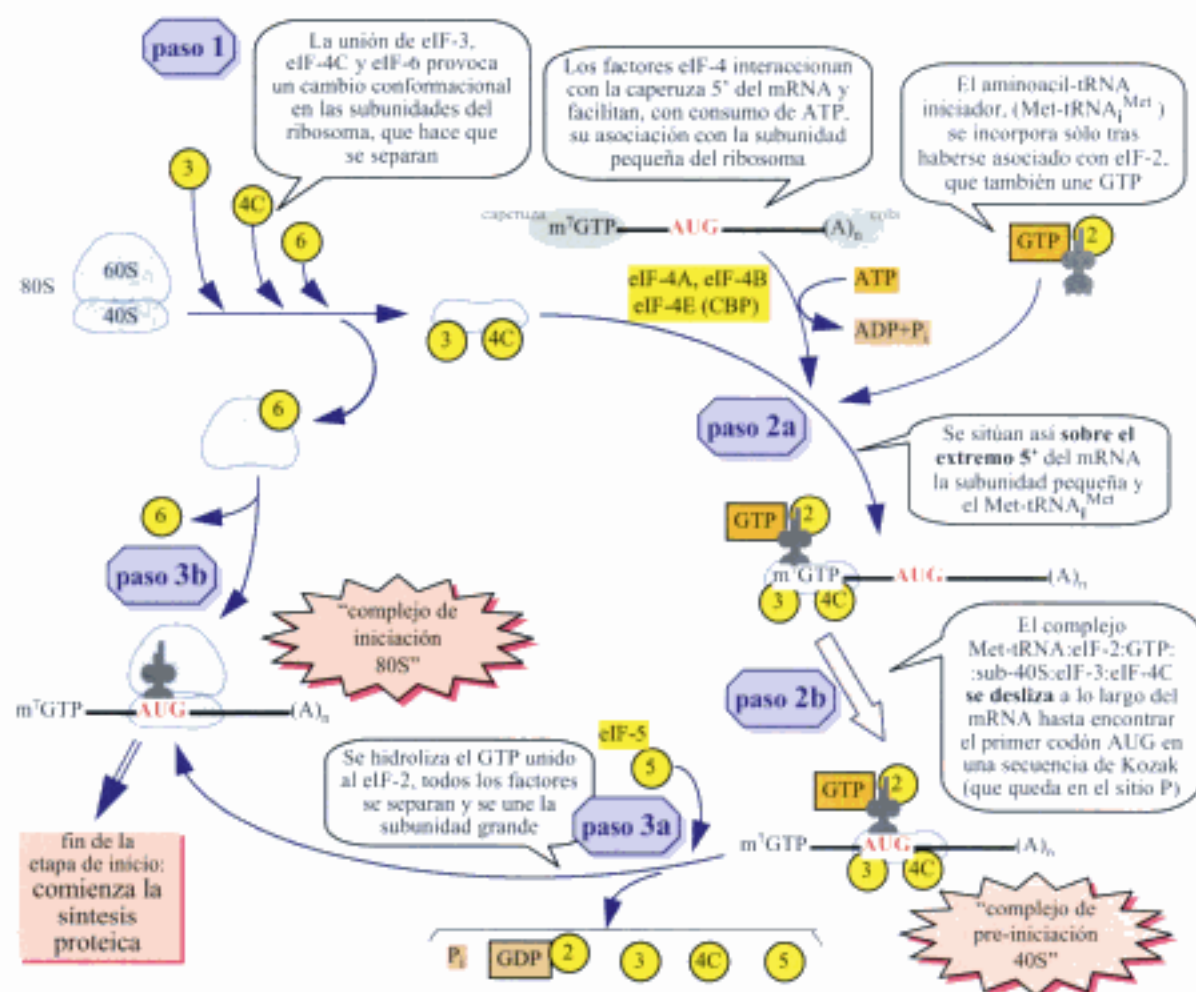
23.3.3 Pasos sucesivos en la fase de inicio

En prácticamente todos estos pasos, así como en las fases de elongación y terminación, intervienen una serie de proteínas llamadas genéricamente **factores proteicos** de iniciación, elongación o terminación. Los **factores de iniciación** se nombran con la raíz **IF** en procariotas, **eIF** en eucariotas y en general **(e)IF**. A partir de este momento, sólo se describe la traducción en eucariotas; para facilitar su seguimiento, se establecerá una división en 9 pasos sucesivos: 3 en la iniciación, 3 en la elongación y 3 en la terminación.

Resumiendo, el inicio de la traducción requiere la unión del tRNA al codón de inicio AUG en el mRNA, en el entorno del **sitio P o peptidilo** del ribosoma. No intervienen en esta fase el codón anejo (corriente abajo) al AUG ni los **sitios A o aminoacilo** y **E o de salida** (*exit*) del ribosoma (pág. 311).

23.3.3.1 Disociación del ribosoma (paso 1)

Debe indicarse que las dos subunidades del ribosoma (40S y 60S en eucariotas), sintetizadas en el núcleo y exportadas a través de los poros nucleares al citosol, se ensamblan en éste formando el ribosoma 80S (véase pág. 80 para detalles de estructura y composición). Para su intervención en la traducción, el ribosoma debe disociarse en sus dos subunidades para permitir, en las subetapas siguientes, la unión de tRNA y mRNA y el inicio de la síntesis.



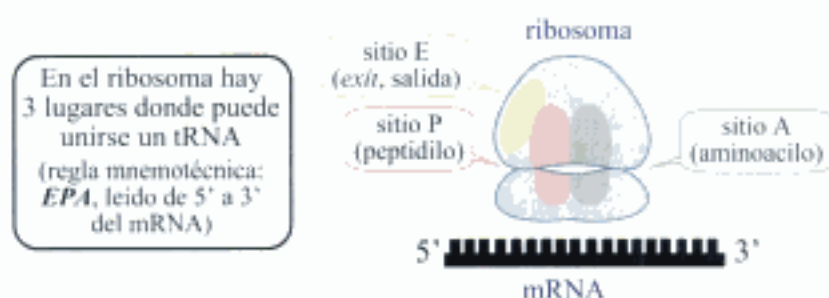
23.4 FASE 2: ELONGACIÓN O ALARGAMIENTO DE LA CADENA PEPTÍDICA

Esta fase incluye desde la adición del segundo aminoácido hasta el último. Se trata de un proceso cíclico que se repite tantas veces como aminoácidos se incorporen. Cada ciclo consta de 3 pasos: ubicación del nuevo aa-tRNA en el sitio A del ribosoma, formación del enlace peptídico y translocación del peptidil-tRNA al sitio P. En ellos intervienen tres factores proteicos citoplasmáticos o **factores de elongación**: **eEF-1 α** , **eEF-1 $\beta\gamma$** y **eEF-2** (los factores procarióticos correspondientes, EF-Tu, EF-Ts y EF-G son similares y actúan de la misma forma; EF-Tu es la proteína más abundante en la célula de *E. coli*).

Debe resaltarse que la etapa de elongación no posee la capacidad de diferenciar el aminoácido que se incorpora, sino que es específica únicamente con respecto al anticodón del aa-tRNA; la incorporación de un aminoácido correcto viene determinada por la actuación previa de la aminoacil-tRNA sintetasa durante la etapa de activación. Esto se demostró al hacer reaccionar Cys-tRNA^{Cys} con hidruro de níquel para eliminar el grupo tiol de Cys y obtener Ala-tRNA^{Cys}. Éste, aunque sigue reconociendo los anticodones correspondientes a Cys según el código, conduce a la incorporación de Ala en las posiciones donde el polipéptido debería tener Cys.

23.4.1 Ubicación del aminoacil-tRNA en el sitio A (paso 4)

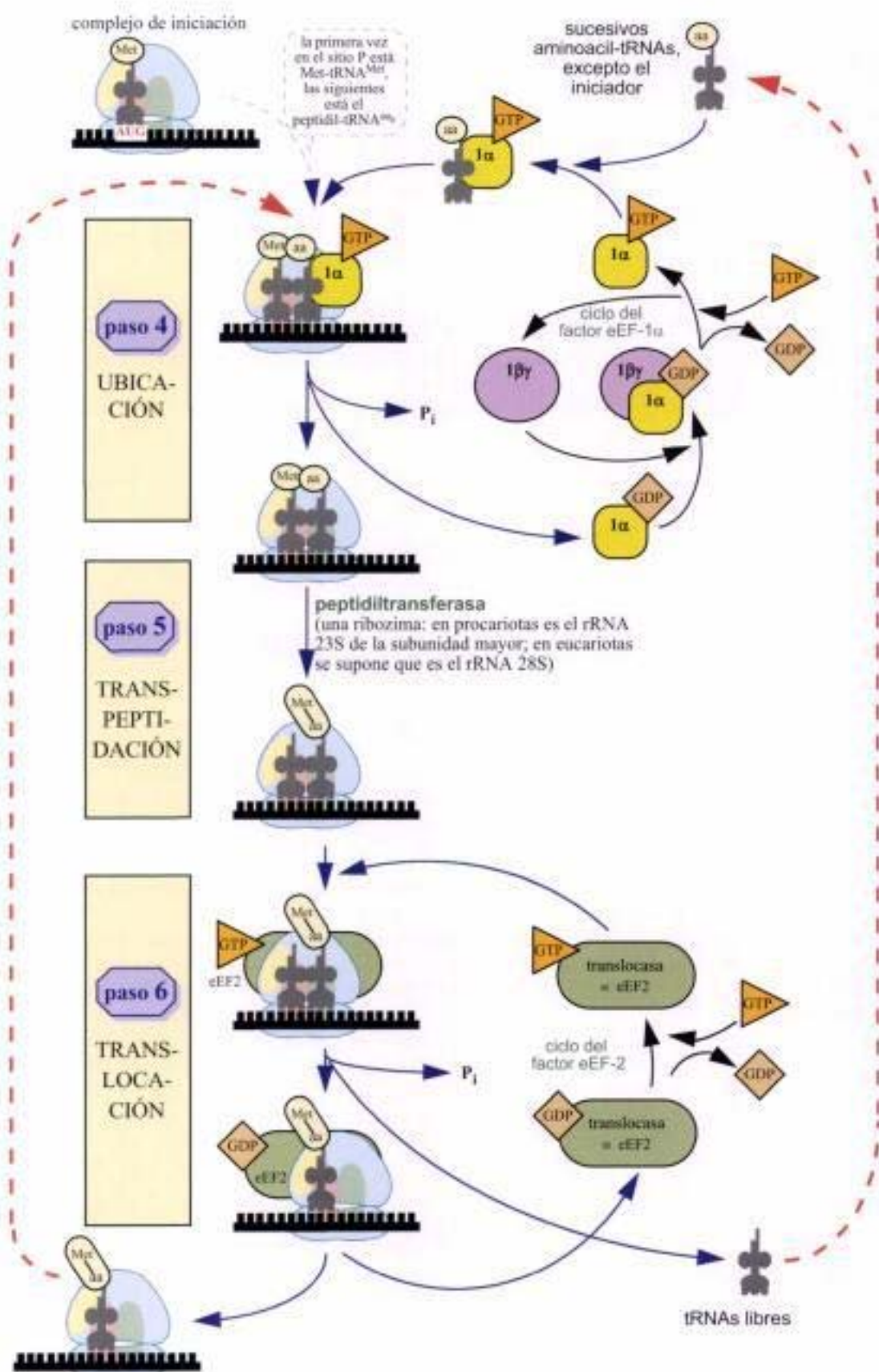
Tras incorporarse el Met-tRNA al sitio P del ribosoma en el complejo de iniciación 80S, queda aún vacío el sitio A. En él se producirá la incorporación de cada nuevo aminoácido (en la primera vuelta del ciclo de elongación, el aminoácido nº 2), por apareamiento del anticodón de su aa-tRNA con el codón correspondiente del mRNA.



La entrada de todos los sucesivos aa-tRNA en el sitio A requiere que estén unidos al factor de elongación **eEF-1 α** que, similarmente a lo ya estudiado para el factor de iniciación eIF-2, une GTP y tiene actividad GTPasa latente. (La estructura de los factores eIF-2, eEF-1 α y eEF2 es similar entre sí y a la de otras proteínas G, lo que sugiere su procedencia de una proteína ancestral común.) De hecho, la forma habitual *in vivo* de casi todos los aa-tRNAs es como complejo ternario aa-tRNA : eEF-1 α : GTP. Son estos complejos los que difunden al sitio A vacío del ribosoma y reconocen en él sitios de unión adecuados. Si el apareamiento entre anticodón y codón es correcto, el aminoacil-tRNA se ubica de forma estable en el ribosoma, lo que provoca cambios conformacionales que estimulan la actividad GTPasa de eEF-1 α . La forma GDP de eEF-1 α no posee afinidad por el ribosoma, liberándose y quedando así el aa-tRNA unido al sitio A, dispuesto para las etapas siguientes.

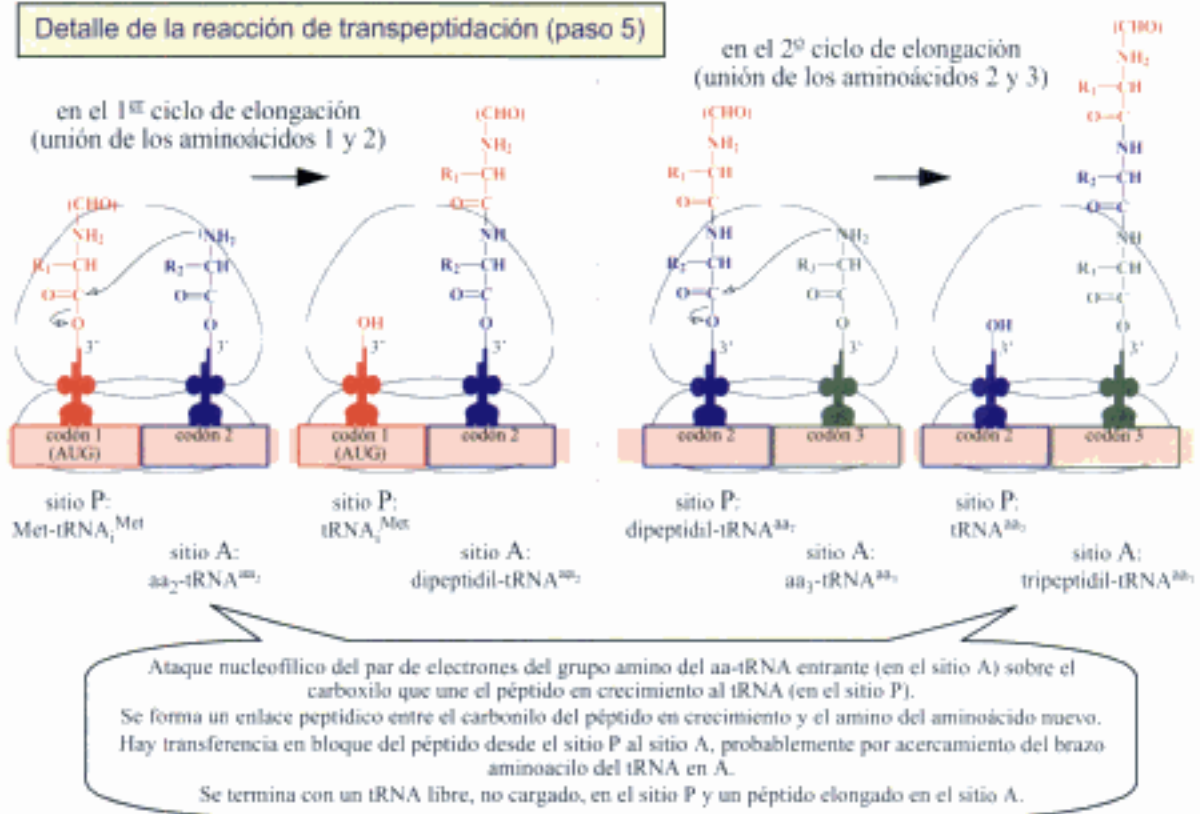
De forma análoga a lo que ocurre con el factor de iniciación eIF-2 (pág. 310), se regenera la forma eEF-1 α : GTP por intercambio de nucleótidos mediado por un nuevo factor de elongación, **eEF-1 $\beta\gamma$** ; se produce, pues, el **cuarto gasto energético** de la traducción (una molécula de ATP por cada aminoácido incorporado). eEF-1 α queda así disponible para incorporar un nuevo aminoácido.

Debe indicarse que eEF-1 α interacciona con cualquier aa-tRNA excepto con el iniciador, Met-tRNA^{Met}, que se distingue de los restantes tRNAs en la estructura secundaria del tallo aceptor. De este modo, durante la elongación, el tRNA iniciador no interviene en la incorporación de Met internas. Esta especificidad, junto con la ya descrita en la fase de inicio, mantienen la dualidad de tRNAs para Met.



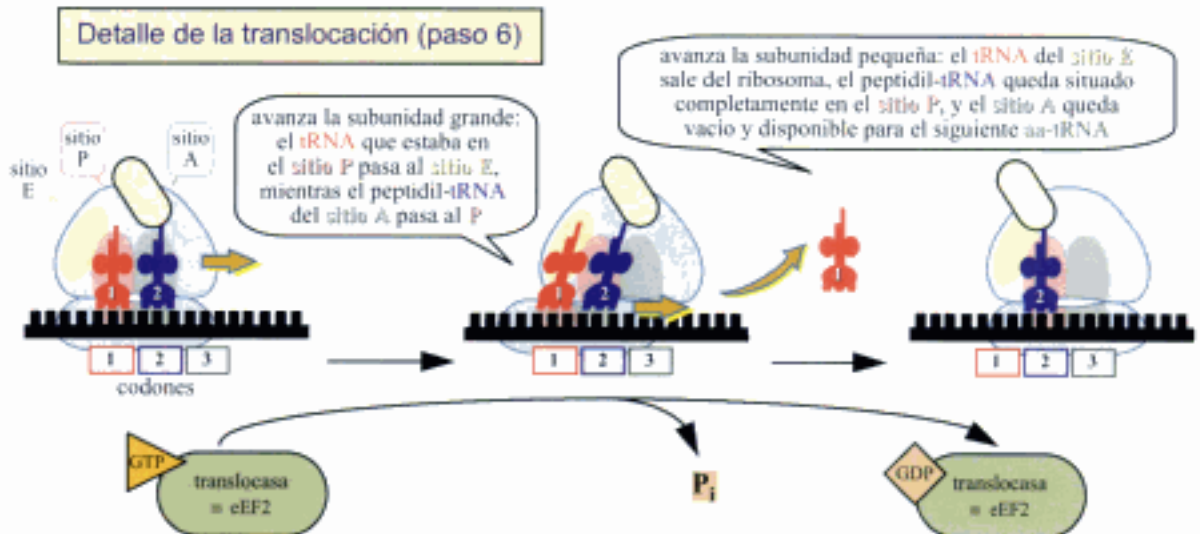
23.4.2 Transpeptidación: formación del enlace peptídico (paso 5)

Una vez situadas las moléculas de aa-tRNA y Met-tRNA^{Met} (en ciclos sucesivos, peptidil-tRNA) en los sitios A y P del ribosoma, respectivamente, se forma el enlace peptídico entre ambos. Esta reacción la cataliza una actividad enzimática peptidiltransferasa (o transpeptidasa), localizada probablemente en el rRNA 28S, componente de la subunidad grande 60S; se trata, pues, de una ribozima (pág. 78), aunque las proteínas ribosómicas contribuyen posiblemente a que sea activa.



23.4.3 Translocación: desplazamiento del ribosoma en un codón (paso 6)

El ribosoma completo se desplaza en sentido 5'→3' del mRNA (corriente abajo), cada vez un solo codón, desde su ubicación inicial sobre AUG. El responsable de este desplazamiento es el tercer factor proteico de elongación, **eEF-2**, conocido como **translocasa**. Este factor interacciona principalmente con la subunidad mayor, originando cambios conformationales en el ribosoma completo que justifican el desplazamiento. Al mismo tiempo, impide la unión de un nuevo aa-tRNA en el sitio A, evitando que se pase al siguiente ciclo antes de tiempo.



En esta etapa tiene lugar el **quinto gasto energético**, pues la energía requerida para cada translocación se obtiene de la hidrólisis de GTP gracias a una actividad GTPasa que reside en el propio factor eEF-2 o en el ribosoma. Tanto en este caso como en los anteriores, la hidrólisis de GTP está acoplada a cambios conformacionales en la maquinaria de traducción, que conducen a que se retire del complejo el factor que une el nucleótido y posee la actividad GTPasa.

23.4.4 Repetición del ciclo de elongación (pasos 4, 5 y 6)

Al terminar la primera translocación se ha completado la primera vuelta o ciclo de elongación, llegando a una situación idéntica a la existente al comienzo: sitio A libre y sitio P ocupado por un tRNA. La diferencia estriba en que el Met-tRNA_i del sitio P ha sido sustituido por un dipeptidil-tRNA. La elongación continúa de forma cíclica (figura en pág. 312) hasta que se añade el último aminoácido. La llegada al sitio A (situado sobre los codones 2, 3, 4, etc.) de otros aa-tRNAs permite la unión de los siguientes aminoácidos en sucesivos ciclos de elongación. Al acabar cada uno de ellos, el péptido en crecimiento siempre queda unido al tRNA del último aminoácido incorporado. Los enlaces peptídicos se forman siempre sobre el grupo NH₂ del último aminoácido incorporado y, por tanto, la cadena polipeptídica crece desde su extremo amino-terminal hacia el carboxilo-terminal.

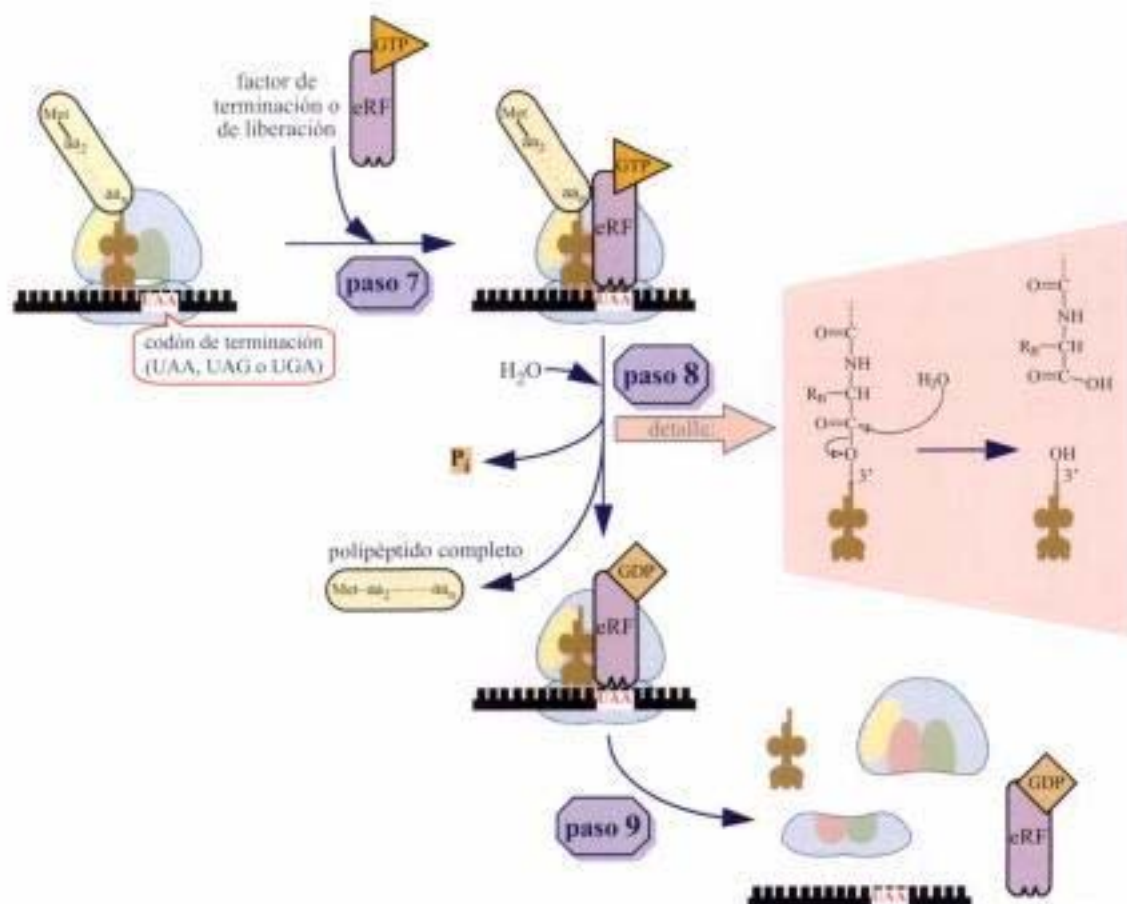
23.5 FASE 3: TERMINACIÓN

Comprende los pasos necesarios para liberar el polipéptido ya completo de su unión al tRNA en el sitio P y, al mismo tiempo, disociar el ribosoma del mRNA. Esto sólo puede ocurrir en respuesta a la presencia en el sitio A de uno de los 3 codones de terminación (UAA, UAG o UGA) en el caso de mRNA codificado por DNA nuclear, y UAA, UAG, AGA o AGG en el caso de mRNA codificado por DNA mitocondrial.

En eucariotas interviene en el proceso un único factor proteico de terminación o de liberación, **eRF** (*Releasing Factor*), en contraste con procariotas donde existen tres (RF1, RF2 y RF3).

23.5.1 Unión del factor de liberación (paso 7)

No existe ningún tRNA que reconozca los codones de paro, por lo que cuando el ribosoma se transloca sobre uno de ellos no puede progresar la elongación. En su lugar, el factor **eRF** se une al ribosoma en el sitio A, frente al codón de terminación.



23.5.2 Hidrólisis del peptidil-tRNA (paso 8)

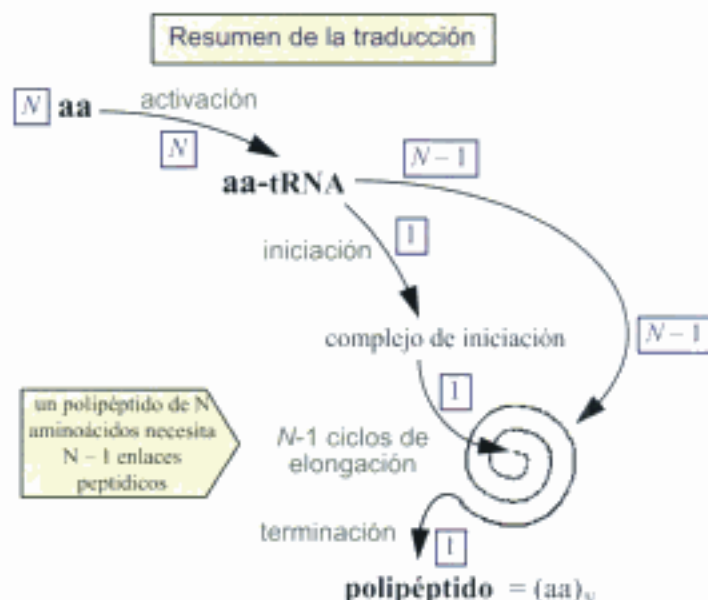
La unión de eRF al ribosoma altera la actividad peptidiltransferasa de tal forma que como agente nucleofílico se emplea el agua (en lugar del grupo amino de un aminoácido, pág. 313). Como consecuencia, se hidroliza el enlace éster del peptidil-tRNA, liberando la cadena polipeptídica. En este proceso además se hidroliza una molécula de GTP asociada al factor eRF, lo que supone el **sexto gasto energético** de la traducción.

23.5.3 Disociación (paso 9)

El tRNA no cargado sale del ribosoma. La forma eRF : GDP ya no posee afinidad por el ribosoma y también se libera. Finalmente, se separan las dos subunidades ribosómicas (en eucariotas, 80S $\rightarrow$ 40S + 60S), liberando el mRNA. Todos los componentes quedan así disponibles para iniciar la síntesis de una nueva proteína (excepto el eRF, que debe recambiar su GDP por GTP, probablemente por un ciclo similar a los estudiados para eIF-2, eEF-1a y eEF-2).

23.6 ENERGÉTICA DE LA SÍNTESIS DE PROTEÍNAS

La síntesis proteica es un proceso energéticamente muy costoso, que utiliza una gran proporción de todo el equivalente en ATP disponible en la célula. Recapitulando todo lo detallado previamente en cada etapa y teniendo en cuenta la repetición cíclica de la elongación se puede calcular el requerimiento energético mínimo (posiblemente la formación o disociación de los numerosos complejos multimoleculares requiera algún aporte energético adicional, así como la corrección de pruebas y, en el caso de procariotas y orgánulos, la formilación).



gasto	fase	paso	(pág.)	nº de enlaces fosfato por paso	nº total de enlaces fosfato por polipéptido de N aminoácidos
nº 1	activación	(aminoacilación del tRNA)	307	2 por cada aa	2N
nº 2	iniciación	nº 2 (incorporación del mRNA)	309	no determinado	al menos 1
nº 3	iniciación	nº 3 (complejo de iniciación)	310	1 por polipéptido	1
nº 4	elongación	nº 4 (ubicación del aa-tRNA)	311	1 por aa añadido	N-1
nº 5	elongación	nº 6 (translocación)	313	1 por aa añadido	N-1
nº 6	terminación	nº 8 (hidrólisis)	315	1 por polipéptido	1
total:					4N + 1 (al menos)

Esta hidrólisis de cuatro enlaces fosfoanhidrido por cada aminoácido representa un gran cambio en energía libre. Concretamente, $-30,5 \times 4 = -122 \text{ kJ/mol de aminoácido}$, muy superior a la energía del enlace peptídico (energía libre estándar de hidrólisis de unos -21 kJ/mol). Aunque pueda parecerlo, esto no es un despilfarro energético, sino que la mayor parte de la energía extra (-101 kJ/mol) sirve para compensar la pérdida de entropía que tiene lugar durante la síntesis proteica. Este descenso de entropía se debe principalmente al ordenamiento de los aminoácidos en el polipéptido. También se pierde entropía cuando un aminoácido se une a un tRNA particular y cuando un aa-tRNA se asocia a un codón específico. Así, el aparente exceso energético se emplea en conseguir la casi perfecta fidelidad en la traducción biológica del mensaje genético del mRNA a la secuencia de aminoácidos de las proteínas.

23.7 INHIBIDORES DE LA TRADUCCIÓN

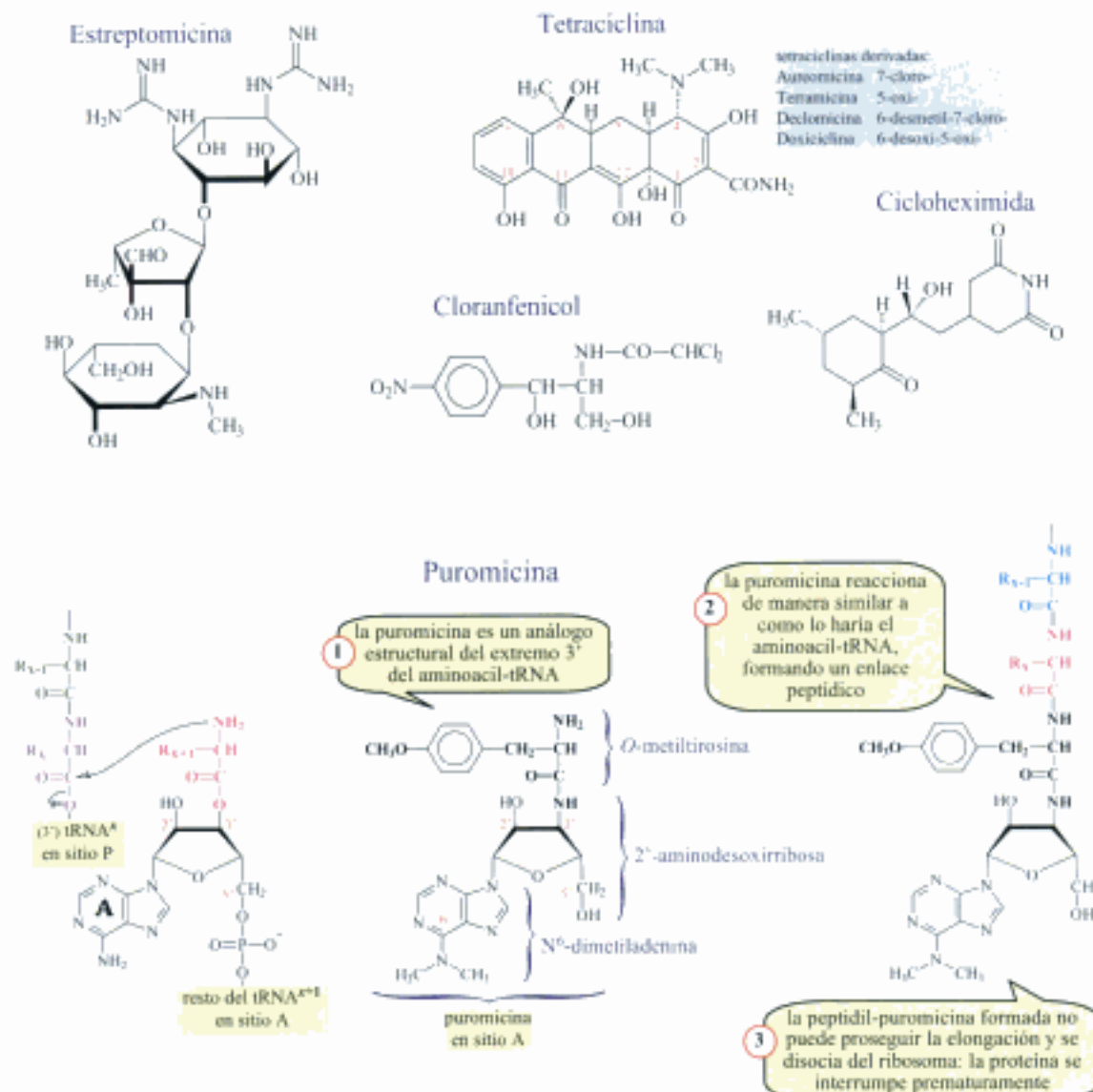
Numerosos y diversos compuestos actúan impidiendo o dificultando la traducción. Su estudio es de interés por varios motivos:

- Desde el punto de vista práctico, porque muchos de ellos se pueden emplear como antibióticos, especialmente cuando ejercen su efecto selectivamente sobre células procariotas (bacterias), afectando poco o nada a las eucarióticas. Precisamente el estudio de sus mecanismos de acción ha contribuido a lo largo de muchos años al esclarecimiento de los detalles moleculares del proceso de traducción. A este respecto, debe mencionarse que el uso indiscriminado (sin control médico) de los antibióticos ha creado una resistencia bacteriana de gran trascendencia tanto con respecto a la flora bacteriana intestinal como a la terapia de las infecciones; estos aspectos deben ser abordados de forma especializada, en tratados de Microbiología.
- Desde un punto de vista teórico, por el papel central de la traducción en la vida celular. Son muchos los antibióticos y toxinas que bloquean o interfieren en ciertos pasos de la síntesis proteica, de forma más o menos específica, en procariotas y eucariotas. Esta especificidad es la mejor prueba de la diferencia entre la maquinaria de traducción en ambos tipos de organismos.

El efecto inhibitor puede ejercerse en distintas etapas del proceso de traducción:

Etapa afectada	Ejemplos	Efecto sobre procariotas o eucariotas	Detalles de la acción
Activación de los aminoácidos	Mupirocina (o ácido pseudomónico)	P	Inhibe competitivamente la Ile-tRNA sintetasa, evitando la incorporación de Ile y deteniendo la síntesis proteica
Iniciación (paso 2, complejo de preiniciación)	Estreptomicina (un aminoglicósido)	P	Se fija de modo irreversible a la subunidad menor 30S, por interacción con varias de sus proteínas y con el rRNA 16S. Distorsiona la entrada de fMet-tRNA iniciador y también produce errores de lectura del mRNA durante la elongación, al interferir con el apareamiento codón/anticodón
	Pactamicina	E	Impide la ubicación de Met-tRNA iniciador en la subunidad menor 40S
	Showdomicina	E	Impide la formación del complejo Met-tRNA _i : eIF-2 : GTP
	Interferón	E	Induce la expresión de una proteína quinasa que fosforila a eIF-2, inactivándolo (de forma similar al HCl, pág. 320)
Iniciación (paso 3, complejo de iniciación)	Eritromicina	P	Se une a un sitio específico en el rRNA 23S de la subunidad mayor 50S. Impide la asociación con la subunidad menor; otra acción posible es interfiriendo la translocación
Elongación (paso 4, ubicación)	Tetraciclina	P > E	Se unen a la subunidad menor, interfiriendo con la fijación del aa-tRNA al sitio A
	Kirromicina	P	Bloquea la disociación de GDP del factor EF-Tu (equivalente al eEF-1α eucariótico), lo que evita su salida del ribosoma
	Ricina (glicoproteína de origen vegetal; con 2 cadenas, A y B; la primera es la toxina, 66 kDa)	E	Se une a proteínas de la subunidad mayor 60S, bloqueando la unión de aa-tRNA : eEF-1α : GTP, y posiblemente también de eEF2 (transpeptidación)
Elongación (paso 5, transpeptidación)	Cloranfenicol	P y mitocondrial	Se une selectivamente a una proteína de la subunidad mayor 50S, interfiriendo en la interacción del aa-tRNA con el centro activo de la peptidiltransferasa, como inhibidor competitivo
	Cicloheximida	E	Similarmente al cloranfenicol, inhibe la actividad peptidiltransferasa pero sólo en el ribosoma eucariótico (subunidad mayor 60S)
	Puromicina	P y E	Análogo estructural del aa-tRNA, forma enlace con el péptido provocando su terminación prematura

Elongación (paso 6, translocación)	Aminoglicósidos distintos de la estreptomicina (ejs.: gentamicina, kanamicina, neomicina)	P	La mayoría interaccionan con el rRNA 16S de la subunidad mayor 50S, impidiendo la fijación del factor de elongación G (EF-G, equivalente al eEF-2 eucariótico). Los mecanismos de acción son diversos
	Esparsomicina	E	Inhibe la translocación
	Ácido fusídico	P y E	Se une de forma muy estable al complejo EF-G : GDP : ribosoma, impidiendo la liberación de EF-G. También inhibe a eEF-2 (factor eucariótico equivalente)
	Toxina diftérica (proenzima proteica de 62 kDa; la toxina es un fragmento de 21 kDa)	E	Cataliza una reacción que inactiva eEF-2 de forma irreversible (ADP-ribosilación a costa de NAD, sobre un residuo His modificado de eEF-2)
Terminación	No se conoce ninguno		



La selectividad de la mayoría de antibióticos y toxinas sobre organismos procariotas o eucariotas es posible gracias a las diferencias en sus ribosomas, así como, en algunos casos (por ejemplo tetraciclinas), la incapacidad de la molécula de atravesar la membrana celular de organismos superiores.

Desde un punto de vista básico, sólo debemos referirnos a una situación general, la de la célula de un individuo adulto. Entre las proteínas específicas que se unen al mRNA recién sintetizado y ayudan a su transporte fuera del núcleo se encuentra un factor proteico multimérico, que incluye al **eIF-4E** o **proteína de unión a la caperuza (CBP)**, ya estudiada por su papel en el inicio de la traducción (pág. 310).

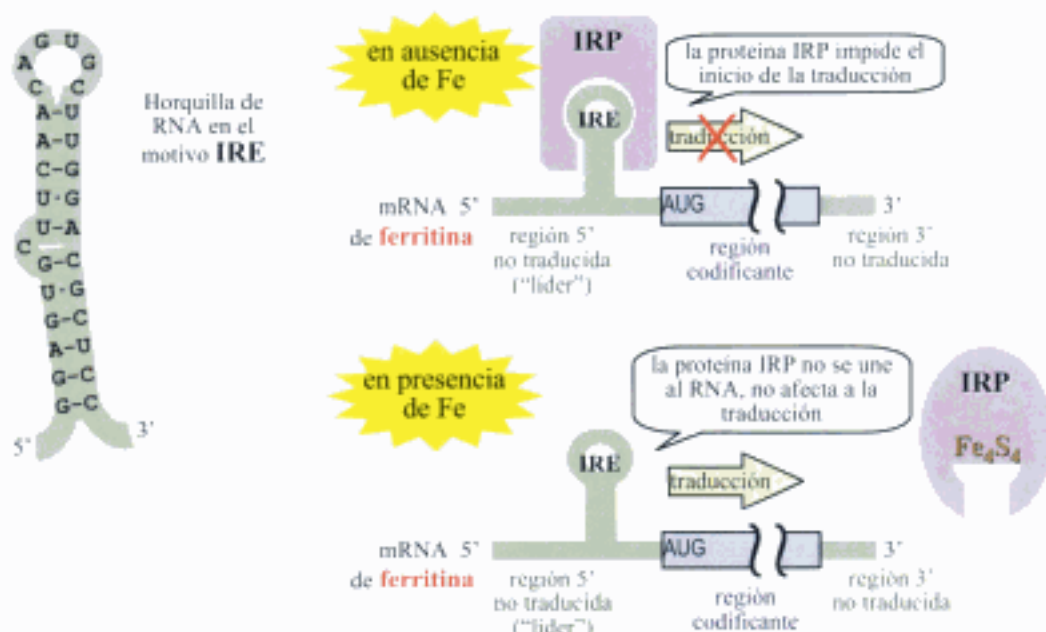
23.8.2 Control de la capacidad y frecuencia de traducción de los mRNAs

Una vez que el mRNA ha alcanzado el citoplasma, el control de su traducción se ejerce esencialmente en la iniciación, haciendo de ésta la etapa limitante. Principalmente se regula la formación del complejo de preiniciación 80S, la actividad de los factores de inicio eIF-2 y eIF-4E, o bien a través de la influencia de estructuras secundarias en el mRNA, por ejemplo regiones en doble hebra intramolecular, que pueden enmascarar los sitios de unión del ribosoma y el codón de inicio. Se estudian a continuación con cierto detalle dos ejemplos cuyo mecanismo es bien conocido.

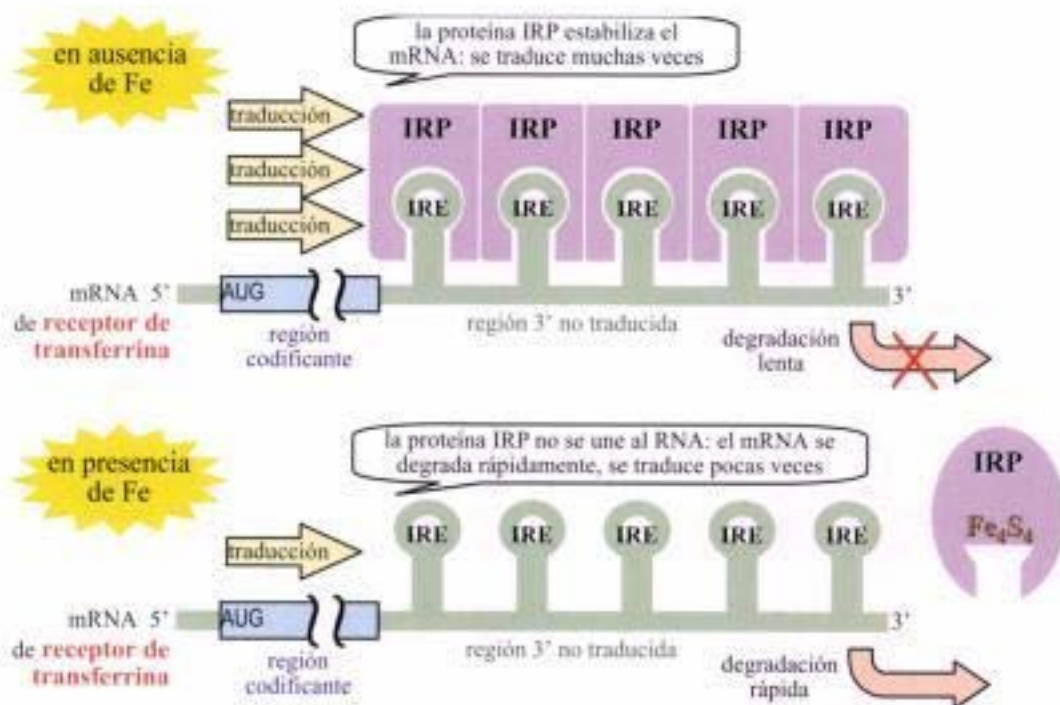
23.8.2.1 Ejemplo 1: Control de la síntesis de ferritina por bloqueo del mRNA

La **ferritina** es una proteína cuya función es almacenar hierro, como reserva para la biosíntesis de hemoglobina, mioglobina, citocromos y otras muchas proteínas, y al mismo tiempo protegiendo a la célula contra la toxicidad del hierro libre. Su síntesis se controla a través de las proteínas **IRP**, que se unen a la región líder 5' no traducida del mRNA impidiendo el inicio de la traducción.

Las **IRP** o **proteínas reguladoras del hierro** (*iron-regulatory proteins*) poseen la capacidad de unirse al mRNA, con distinta afinidad en función de la concentración intracelular de ese metal. Por ejemplo, una de ellas, la **IRP-1**, cuando el hierro es abundante lo incorpora en forma de núcleo hierro-azufre (*cluster* Fe_4S_4) y no interacciona con el mRNA. Por el contrario, cuando el hierro es escaso en la célula la **IRP-1** no presenta el núcleo Fe-S y adopta una conformación que le permite unirse con elevada afinidad a una horquilla de RNA con bucle (pág. 59). Ésta se conoce como **motivo IRE** o **elemento de respuesta a hierro** (*iron-response element*), por lo que las **IRP** se llaman también **proteínas ligantes de IRE** (**IRE-BP**, *IRE-binding proteins*). Existe un motivo IRE en la región 5' no traducida (pág. 245) del mRNA de ferritina, al que se unen las **IRP** impidiendo el inicio de la traducción y reduciendo así la concentración intracelular de ferritina. La célula se adapta así a la escasez de hierro, evitando la síntesis de una proteína que no es necesaria. Por el contrario, en condiciones de exceso de hierro, se sintetiza libremente la ferritina, para poder fijarlo.

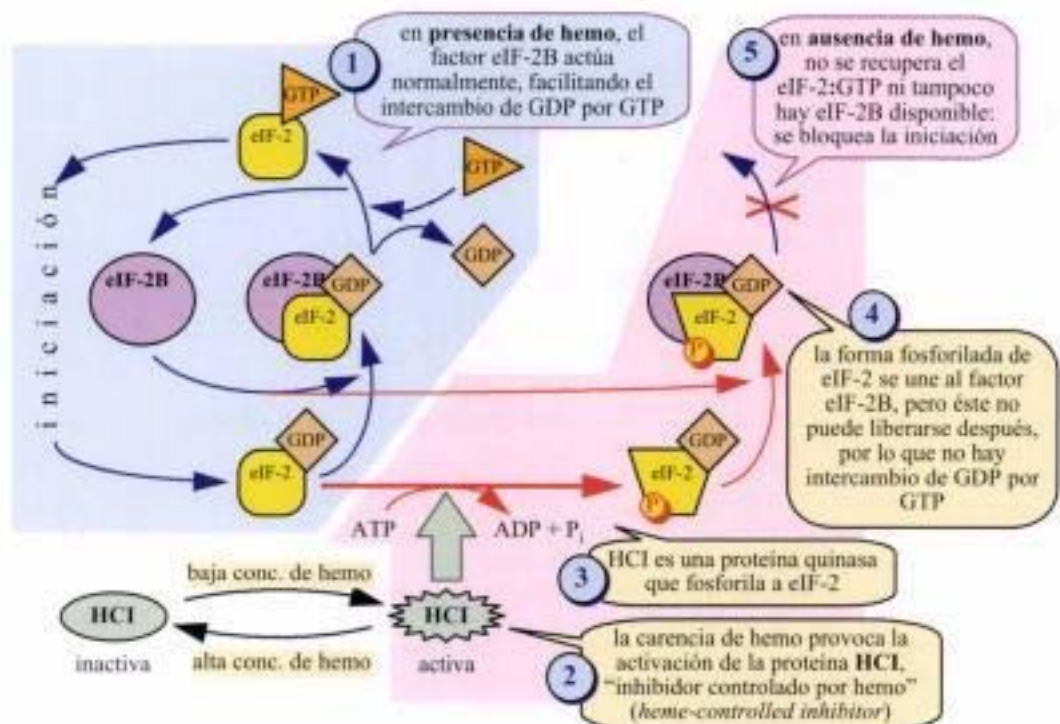


Existe un ejemplo muy similar de regulación, coordinada con la anterior, pero en este caso controlando la vida media del mRNA, lo que corresponde a una regulación **pretraduccional** (pág. 286). Se trata del **receptor de transferrina**, una proteína de membrana necesaria para que las células capten el hierro. En condiciones de escasez de hierro la proteína **IRP** se une también a este mRNA, que posee 5 motivos IRE, pero al estar éstos en su región 3' no traducida el efecto es de protección del mRNA frente a la degradación. En consecuencia, se sintetizan más moléculas del receptor de transferrina a partir de un mismo mRNA, capacitando a la célula para captar hierro con más eficacia. En la situación contraria (abundancia de hierro) **IRP** no es activa para unirse al mRNA, que se degrada más rápidamente, reduciendo la síntesis del receptor, ya no tan necesario. Ambos mecanismos de regulación (traducción del mRNA de ferritina y degradación del mRNA del receptor de transferrina) actúan al unisono con un mismo objetivo, la adaptación de la célula a la disponibilidad de hierro.



23.8.2.2 Ejemplo 2: Regulación de la síntesis de hemoglobina por fosforilación de un factor de iniciación

La síntesis de hemoglobina requiere cantidades estequiométricas de hemo y de globinas. La síntesis de las globinas tiene lugar en precursores eritroides y en reticulocitos (precursores inmediatos de los eritrocitos), donde se expresan activamente los genes correspondientes. La velocidad de traducción del mRNA de globina está regulada en función de la concentración de hemo en la célula, de modo que si ésta descende, disminuye la síntesis. La célula evita así la producción inútil de la proteína para la que no hay grupo prostético (hemo) disponible. La regulación tiene lugar por inhibición del inicio de la traducción, interfiriendo en la reutilización del factor eIF-2 (pág. 310).



Debe hacerse notar que resulta inhibida toda nueva incorporación de una Met inicial: puede continuar la síntesis de las cadenas ya iniciadas, pero no puede iniciarse una nueva. Por otra parte, a diferencia de los ejemplos anteriores (ferritina y receptor de transferrina), el control no afecta específicamente al mRNA de una proteína, sino a la traducción de todos los mRNAs de la célula. Este aparente perjuicio potencial para la célula está justificado porque se trata de células muy especializadas, que sintetizan casi exclusivamente globina.

Tema 24

Modificaciones postraduccionales: maduración, plegamiento, tráfico y degradación de proteínas

24.1	Introducción	322
24.2	Tráfico o destino de las proteínas.....	322
24.2.1	Compartimentos implicados.....	322
24.2.2	Tráfico de proteínas en el citosol y entre orgánulos	323
24.2.3	Tráfico de proteínas de secreción.....	324
24.2.3.1	Características del péptido señal.....	324
24.2.3.2	Mecanismo de entrada de la proteína al lumen del RE.....	324
24.3	Maduración o procesamiento del polipéptido nascente.....	326
24.3.1	Maduración por modificación de aminoácidos.....	326
24.3.1.1	Unión de sustituyentes a los aminoácidos.....	326
a)	Acetilación.....	326
b)	Carboxilación.....	327
c)	Fosforilación.....	327
d)	Hidroxilación.....	328
e)	Metilación.....	328
24.3.1.2	Incorporación de glúcidos: glicosilación.....	328
a)	Características de las glicoproteínas.....	328
b)	Tipos de glicosilación.....	329
c)	Biosíntesis de las glicoproteínas.....	332
24.3.1.3	Modificación con lípidos.....	334
a)	Acilación (ácidos grasos).....	334
b)	Prenilación (terpenos).....	335
24.3.1.4	Formación de puentes disulfuro.....	335
24.3.1.5	Otras modificaciones.....	336
24.3.2	Maduración por escisión de las proteínas: Procesamiento proteolítico.....	336
24.3.2.1	Activación de proenzimas proteolíticas.....	336
24.3.2.2	Activación de precursores de hormonas peptídicas.....	337
24.4	Plegamiento de proteínas.....	337
24.4.1	Mecanismo del plegamiento de las proteínas.....	337
24.4.2	Proteínas implicadas en el plegamiento: carabinas moleculares.....	338
24.4.2.1	Tipos de carabinas.....	339
a)	Carabinas del tipo Hsp70.....	339
b)	Carabinas del tipo Hsp60.....	339
24.4.2.2	Función de las carabinas.....	340
24.5	Degradación de las proteínas.....	340
24.5.1	Degradación lisosómica.....	341
24.5.2	Degradación citosólica.....	341
24.5.2.1	Ubiquitina.....	341
24.5.2.2	Señales desencadenantes de la ubiquitinación.....	342
a)	Residuo amino-terminal.....	342
b)	Secuencias PEST.....	342
24.5.2.3	Proteasas responsables de la degradación.....	342

24.1 INTRODUCCIÓN

La síntesis proteica no termina con la liberación del polipéptido del complejo de traducción, sino que en todos los organismos se requiere una serie de procesos para que el mensaje lineal o unidimensional del polipéptido (secuencia de aminoácidos determinada por la secuencia de nucleótidos del mRNA) se transforme en un mensaje tridimensional, la estructura nativa de la proteína responsable de su función. Estos procesos, recogidos de forma conjunta bajo el término de **maduración o modificación postraduccional**, constituyen el colofón de la síntesis proteica, que a veces se considera la 5ª fase de la traducción, después de la activación del aminoácido, iniciación, elongación y terminación. Por diversas razones, especialmente por un estudio escaso o parcial hasta hace pocos años, estos procesos no han recibido la consideración que merecen en cuanto a su importancia para la célula. Los avances logrados ya permiten tener una visión global, aunque aproximada, del recorrido que experimenta la proteína desde su síntesis hasta su función.

La maduración postraduccional es esencial en todo momento y circunstancia de la célula eucariótica. Cada compartimento subcelular requiere proteínas diferentes, que no se sintetizan en él, para sus variadas funciones: actividad catalítica de reacciones en todas las vías metabólicas, papel estructural en la membrana celular, transporte o almacén de moléculas e iones, función motora muscular, transmisión de señales intra- e intercelular, regulación de la expresión génica, etc. Incluso las mitocondrias y cloroplastos, que poseen un genoma propio, necesitan multitud de proteínas codificadas en el genoma nuclear, que luego deben llegar hasta estos orgánulos.

La variabilidad en cuanto a condiciones y circunstancias particulares en cada tipo celular y la amplitud y complejidad de los procesos implicados en la modificación postraduccional hacen difícil su descripción. A este fin, se seguirá un esquema relativamente común y adaptado a la realidad celular, *aparentemente* lógico, dividiendo el procesamiento postraduccional en los siguientes apartados. Debe señalarse, sin embargo, que los 3 primeros procesos tienen lugar generalmente de forma simultánea.

1. Tráfico o destino de las proteínas hacia diferentes localizaciones, subcelulares o extracelulares, para el ejercicio de su función.
2. Maduración o procesamiento del polipéptido. Es el aspecto esencial de la modificación postraduccional. Comprende modificaciones químicas de los aminoácidos y eliminación de fragmentos del polipéptido.
3. Plegamiento correcto del polipéptido hasta alcanzar su conformación biológica activa. Requiere, en muchos casos, una maduración previa por modificación química, con lo que plegamiento y maduración están íntimamente ligados.
4. Es entonces, una vez que la proteína ha llegado a su destino y cumplido su misión, cuando se debe describir también su catabolismo, su degradación hasta aminoácidos.

Con este proceso se cierra definitivamente el "ciclo biológico de la proteína", es decir:



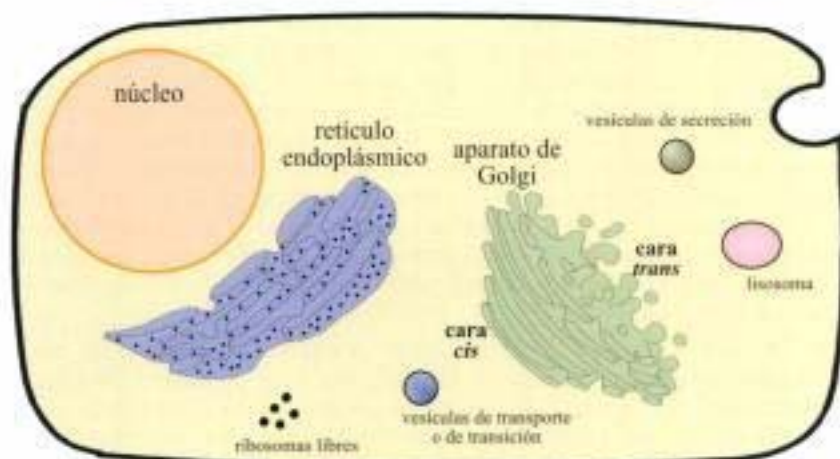
24.2 TRÁFICO O DESTINO DE LAS PROTEÍNAS

Se conoce como tráfico, destino, topogénesis e incluso "clasificación" de proteínas la ruta que siguen estas moléculas en la célula eucariótica hasta alcanzar su localización (intracelular o extracelular) donde pueden ejercer su función, y los mecanismos moleculares responsables de que sigan esa ruta. El tráfico se desarrolla de forma estrechamente conectada con la maduración y el plegamiento del polipéptido.

24.2.1 Compartimentos implicados

El tráfico de proteínas tiene lugar con referencia al citosol, a la membrana plasmática y a los compartimentos cerrados presentes en aquél, los orgánulos delimitados por membranas, constituyentes del llamado "sistema endomembranoso":

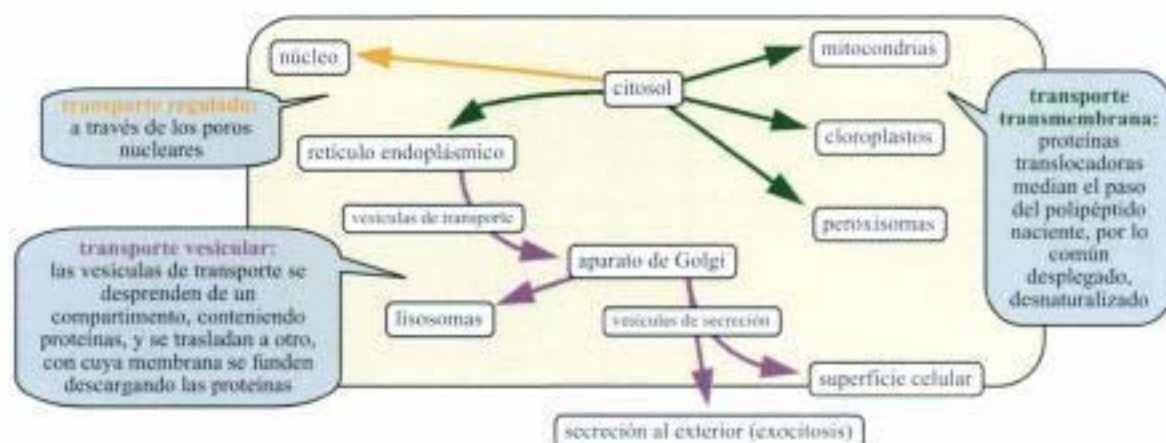
- **Retículo endoplásmico (RE).** Es un sistema intracelular muy ramificado, plegado e interconectado, de estructura variada (canales, esferas, tubos, cisternas aplanadas), con un espacio interior (luz o lumen) aislado del citosol. En su mayor parte, el RE posee ribosomas asociados a su superficie citosólica (formando el **RE rugoso**), donde se realiza la síntesis de las proteínas integrales de membrana (celular, del RE, del Golgi y de lisosomas), de las destinadas al lumen de RE, Golgi o lisosomas y de las destinadas a secreción. Las proteínas con destino y función citosólicos, nucleares o de otros orgánulos (mitocondrias, cloroplastos, peroxisomas...) se sintetizan en ribosomas no asociados al RE. Las proteínas se incorporan al lumen del RE de forma cotraduccional (a la vez que se sintetizan), para allí sufrir los primeros pasos de la maduración. La salida del RE se hace bajo la forma de pequeñas vesículas, que se dirigen al aparato de Golgi, a lisosomas o a la membrana externa.



- **Complejo o aparato de Golgi.** Se trata de un compartimento situado entre el RE y la membrana plasmática en una disposición (cara "cis" frente al RE y cara "trans" frente a la membrana) que facilita la llegada de las vesículas del RE cargadas de proteínas. Está formado por un apilamiento de sacos o bolsas aplanadas, con forma de disco (*dictiosomas*), no interconectadas entre sí. Los polipéptidos sufren aquí nuevos procesamiento hasta dar la proteína madura, ya dotada de las señales moleculares necesarias para definir su destino final. Las proteínas son reenviadas al citosol en nuevas vesículas que se dirigen al orgánulo adecuado o hacia la membrana externa (durante este transporte también se procesan muchas proteínas por escisión proteolítica). En el caso de las proteínas que la célula secreta al exterior, se habla de *vesículas de secreción*, que finalmente se funden con la membrana citoplasmática liberando su contenido al exterior celular por **exocitosis**. Además de esta función, el aparato de Golgi también da lugar a los lisosomas, cargados de enzimas hidrolíticas, que no sólo degradan moléculas intracelulares, sino también partículas extracelulares incorporadas por **endocitosis** (mediante la fusión de endosomas con lisosomas).
- **Mitocondrias.** Son sitios donde tiene lugar específicamente la fosforilación oxidativa, con la producción de ATP para todas las necesidades energéticas de la célula. Sólo unas pocas de las proteínas necesarias están codificadas en el genoma mitocondrial y se sintetizan en ribosomas mitocondriales, por lo que se requiere la entrada a través de su doble membrana de gran cantidad de proteínas sintetizadas en el citosol.
- **Peroxisomas.** Son orgánulos de tamaño pequeño delimitados por una membrana única, ricos en enzimas que intervienen en reacciones oxidativas de moléculas tóxicas.
- **Otros compartimentos.** Quizá con menor trascendencia en lo que respecta al tráfico y la maduración, pero obviamente con importancia funcional y que también son destino específico de sus proteínas respectivas, deben citarse los cloroplastos; lisosomas, etc.

24.2.2 Tráfico de proteínas en el citosol y entre orgánulos

Todas las proteínas comienzan su síntesis en el citosol (excepto las codificadas en los genomas mitocondrial y cloroplástico). La clave del proceso por el que las proteínas se dirigen a su localización son **señales de clasificación** o **secuencias señal** que forman parte de la proteína nascente y dirigen el tráfico de la molécula hacia distintos compartimentos de la célula. Las proteínas que no presentan secuencias señal permanecen en el citosol. Estas señales suelen ser regiones continuas de la secuencia de aminoácidos, de 15 a 60 residuos; frecuentemente situadas cerca de los extremos de la cadena y casi siempre eliminadas por escisión proteolítica (pág. 336) una vez han cumplido su función. En otros casos, sin embargo, la señal es una **"región señal"**, una disposición tridimensional específica, dependiente del plegamiento de la proteína, formada por aminoácidos no contiguos en la estructura primaria.



El aspecto más destacable del tráfico intracelular y más comúnmente estudiado es el que corresponde a las proteínas de secreción. Éste ocurre, de forma muy compleja y versátil, a través de la membrana del RE y complejo de Golgi. Este sistema de transporte implica la introducción de la cadena polipeptídica (sintetizada en la superficie del RE rugoso) en el lumen, a través de la membrana, la salida de la proteína completada en vesículas hacia el complejo de Golgi y, finalmente, su transporte en vesículas de secreción hasta la membrana plasmática y su secreción al exterior. Éste será el único ejemplo que se estudia en detalle, a continuación; no se abordarán otros aspectos concretos, de gran interés pero que escapan al carácter general del presente libro. Entre éstos, pueden citarse la integración de proteínas en la membrana de cualquier orgánulo, el tráfico a orgánulos especiales como mitocondria, cloroplasto o peroxisomas, el transporte bajo la forma de vesículas de secreción desde el RE al complejo de Golgi y a la membrana plasmática, la liberación final de su contenido por exocitosis al exterior celular o la importación de proteínas del medio circundante mediante endocitosis facilitada por receptor.

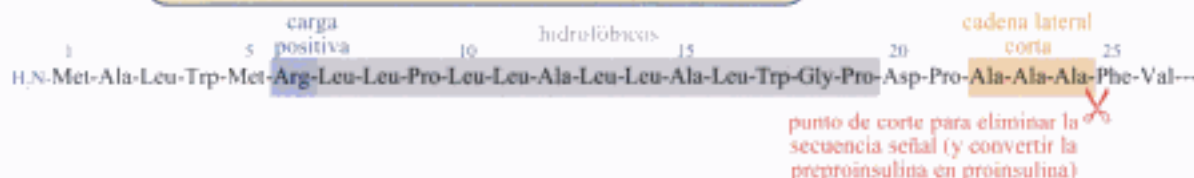
Como ejemplo de otro tipo de destino, cabe mencionar el tráfico de proteínas del citosol a núcleo para su incorporación a la cromatina o su intervención en los procesos de replicación, transcripción y maduración postranscripcional. Estas proteínas deben pasar a través de los *poros nucleares* (estructuras muy complejas, formadas por hasta 100 proteínas diferentes, que atraviesan la doble *envoltura nuclear*). A través de los poros tiene lugar igualmente el tráfico, en sentido inverso, de las moléculas de RNA y las subunidades ribosomales, sintetizadas o ensambladas en el núcleo, respectivamente. Aunque el paso puede realizarse por libre difusión, parece existir una difusión selectiva, sin despliegue de la estructura tridimensional (a diferencia de los otros tipos de tráfico). Se ha propuesto la presencia en las proteínas que se han de transportar al núcleo de una secuencia señal (secuencia de localización nuclear) a través de la cual se unen a otras "proteínas receptores de importación nuclear", que dirigen la entrada a través de los poros, con la energía liberada por la hidrólisis del GTP.

24.2.3 Tráfico de proteínas de secreción

24.2.3.1 Características del péptido señal

La secuencia señal para las proteínas de secreción es normalmente una secuencia peptídica corta en el extremo N-terminal del polipéptido recién sintetizado, que se denomina **secuencia señal o de etiquetado, etiqueta, péptido señal o secuencia líder** (por su posición N-terminal). Presenta uno o más residuos con carga positiva (Lys o Arg), seguidos de 5 a 15 aminoácidos hidrofóbicos y de unos pocos residuos relativamente polares y de cadena corta, especialmente Gly y Ala.

Ejemplo de péptido señal en la preproinsulina humana



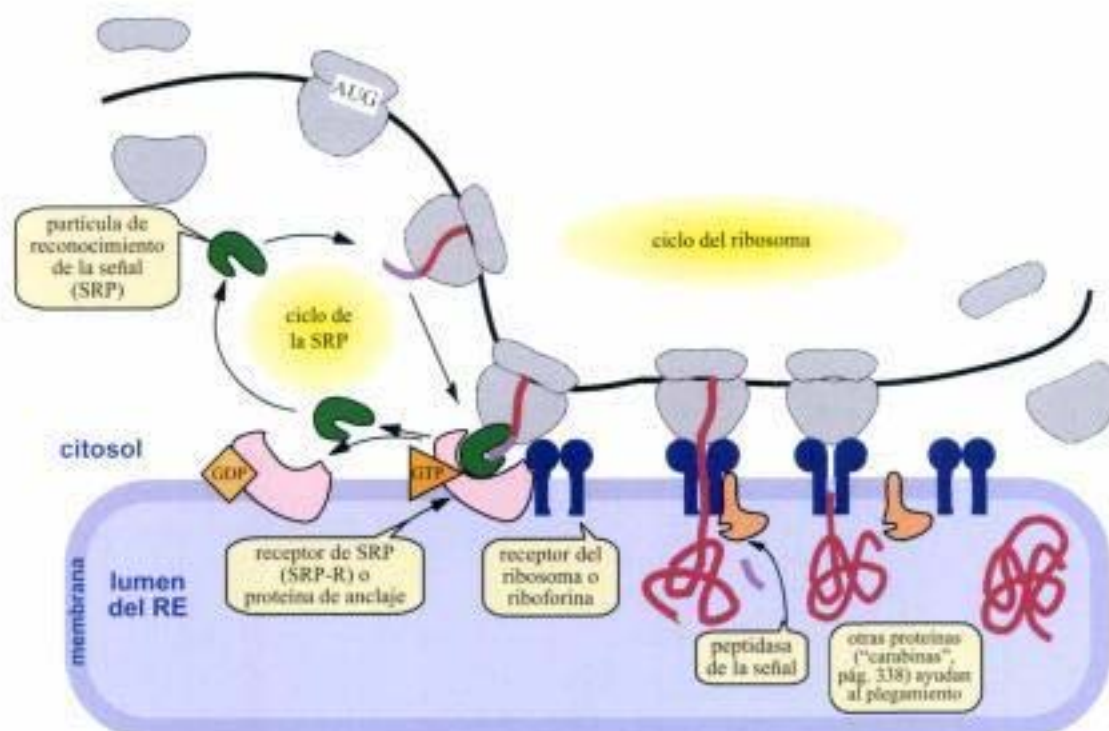
Como se verá a continuación, es frecuente que el péptido señal se elimine, por la acción de una **proteasa de la señal** (o proteasa señal). En estos casos, las proteínas recién sintetizadas, cuya secuencia señal aún no se ha eliminado, se denominan formas **pre-proteína**, para diferenciarlas de la molécula que ya ha sufrido la proteólisis del péptido señal; en algunos casos ésta a su vez es una forma inactiva que precisa un procesamiento adicional por proteólisis, por lo que se la llama **pro-proteína**, y su precursor inicial es, pues, una **pre-pro-proteína** (pág. 336).

24.2.3.2 Mecanismo de entrada de la proteína al lumen del RE

Como se ha indicado, todas las proteínas cuyo destino es la membrana celular, la membrana de RE o Golgi, el interior de estos orgánulos o la secreción al exterior se sintetizan en ribosomas asociados al RE. Además, recién comenzada su síntesis la cadena polipeptídica se introduce al interior del RE, se **transloca** al lumen. La condición única e imprescindible para esta translocación es la presencia en la cadena polipeptídica de la secuencia señal. A continuación se estudia con detalle este proceso, concretándolo para el caso de proteínas de secreción.

1. Primer proceso de unión: La secuencia señal se sintetiza pronto tras el comienzo de la traducción, por estar situada en el extremo amino del polipéptido nascente, y emerge del ribosoma cuando el polipéptido alcanza unos 70 aminoácidos de longitud. La secuencia señal es entonces reconocida rápidamente por un complejo proteico citosólico de gran tamaño (325 kDa), llamado **partícula de reconocimiento de la señal (SRP)**.

Estructuralmente, la SRP es una ribonucleoproteína formada por 6 polipéptidos y un RNA (del grupo de los RNAs citoplásmicos pequeños o scRNAs, llamado en concreto RNA 7SL, con una secuencia relacionada con la familia *Alu* de DNA repetitivo, pág. 114). El RNA 7SL es el responsable directo del reconocimiento de la secuencia señal.



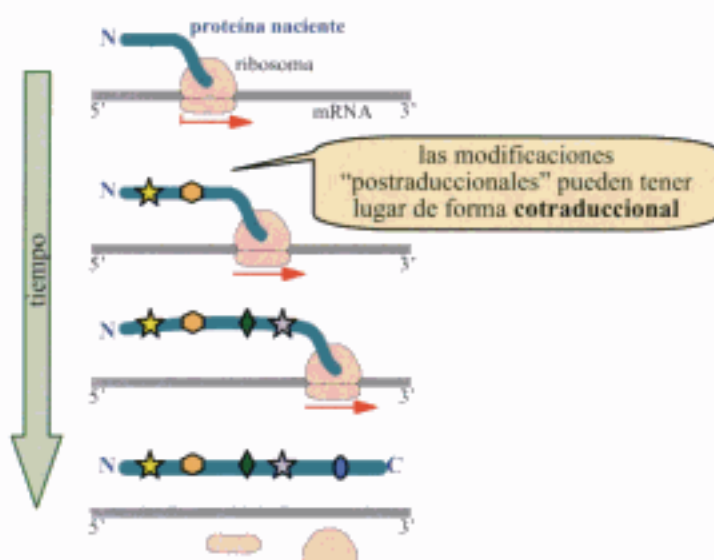
- Segundo proceso de unión: La unión de la SRP provoca una detención temporal de la elongación del polipéptido en el ribosoma, hasta que el complejo anterior (mRNA : ribosoma : aa-tRNA : polipéptido nascente : SRP) contacta con la superficie citosólica del RE, donde la SRP es reconocida y se une a una proteína integral de la membrana del RE, llamada **proteína receptora de SRP (SRP-R) o proteína de anclaje**. Como resultado, todo el complejo queda fijado fuertemente al RE.

El receptor de SRP es una proteína heterodimérica con la capacidad de unir GTP y de catalizar su hidrólisis (actividad GTPasa). La forma con GTP unido es la que interacciona con la SRP, mientras que al hidrolizarlo a GDP pierde afinidad y se libera la SRP.

- Tercer proceso de unión: El complejo ha quedado dispuesto sobre el RE de tal forma que se favorece su interacción con otra proteína integral de la membrana del RE, situada en una posición vecinal a la SRP-R. Esta nueva proteína se llama **proteína receptora del ribosoma o riboforina** porque su misión es, precisamente, la de reconocer la subunidad mayor del ribosoma que forma parte del gran complejo y mediar en la entrada del polipéptido nascente al interior del RE, efectuando un papel como "poro" de la membrana.
 - Como consecuencia de este "anclaje" ocurre lo siguiente: La subunidad α de SRP-R hidroliza el GTP unido, dando GDP y P_i (gasto energético de un GTP por polipéptido).
 - Se disocia la SRP, quedando disponible en el citosol para volverse a emplear ("ciclo de la SRP").
 - Al liberarse la SRP se elimina la inhibición de la traducción, es decir, se reanuda la síntesis proteica, ya en el lumen del RE.
 - En la riboforina se abre un "canal de translocación" que permite la entrada en el lumen de la cadena polipeptídica creciente. De ahí los nombres alternativos de **translocasa** para la riboforina y de **complejo de translocación** para el complejo que se formó.
- En la cara luminal del RE, asociada a la porción interior de las riboforinas del poro, se encuentra una nueva proteína integral de membrana llamada **peptidasa de la señal**, pues escinde el péptido señal, separándolo del resto del polipéptido en crecimiento. Una vez liberado, aquél se degrada hasta aminoácidos.
- La finalización de la síntesis en el lumen del RE, con liberación de la proteína, ocurre cuando el extremo carboxilo del polipéptido ha pasado a través de la membrana del RE. El ribosoma se disocia entonces del RE y se desorganiza todo el complejo de traducción para volver a reciclarse ("ciclo del ribosoma").

24.3 MADURACIÓN O PROCESAMIENTO DEL POLIPÉPTIDO NACIENTE

Conceptualmente es el componente más sencillo de la modificación postraduccional. Puede tener lugar tras haber finalizado la síntesis, una vez liberado el polipéptido del ribosoma o, más comúnmente, de forma simultánea con la traducción, es decir, la región de la molécula que se ha sintetizado antes (N-terminal) va siendo modificada mientras aún se está elongando el polipéptido por una región más próxima al extremo C-terminal. Esta segunda posibilidad recibe el nombre de **modificación cotraduccional**, aunque los procesos sufridos son los mismos en ambos casos. Se utilizan enzimas y cofactores específicos para una multitud de reacciones de modificación.



La maduración implica dos grandes tipos de modificaciones del polipéptido: la transformación de las cadenas laterales de algunos aminoácidos y la escisión proteolítica.

24.3.1 Maduración por modificación de aminoácidos

Diversos aminoácidos del polipéptido, siempre en posiciones específicas, sufren reacciones de modificación química, aunque catalizada enzimáticamente. Se describen a continuación algunas de las más frecuentes.

24.3.1.1 Unión de sustituyentes a los aminoácidos

Es frecuente la adición de determinados grupos (como acetilo, carboxilo, fosfato, hidroxilo o metilo) a la cadena lateral de aminoácidos o a los extremos amino- o carboxi-terminales. Se han descrito más de un centenar de residuos aminoácidos modificados; veamos algunos de los ejemplos más comunes.

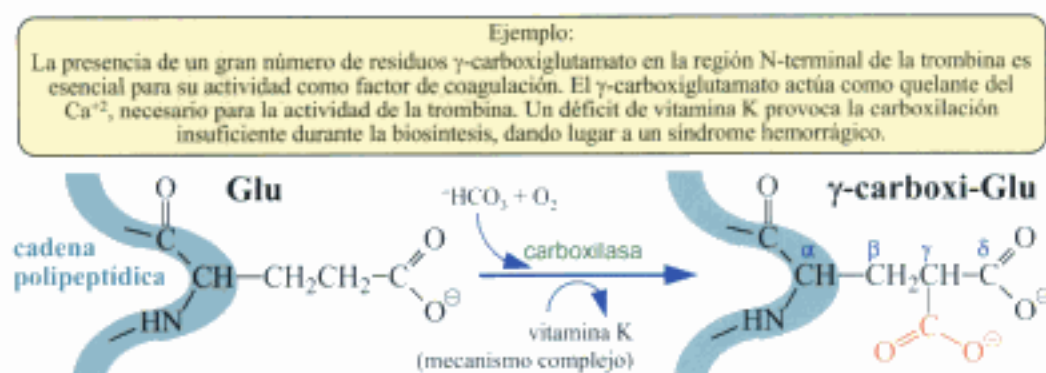
a) Acetilación

La adición de grupos acetilo tiene lugar en un 50% de las proteínas eucarióticas. Una posición frecuente es el grupo **amino terminal**, previa eliminación del residuo de metionina N-terminal (pág. 336). Uno de los efectos de esta modificación es un aumento de la resistencia a la degradación, al proteger el grupo amino. También se acetila el grupo amino en la **cadena lateral** de residuos de lisina en las histonas, de forma reversible; esta modificación es en parte responsable de la regulación de la condensación de la cromatina (pág. 264).



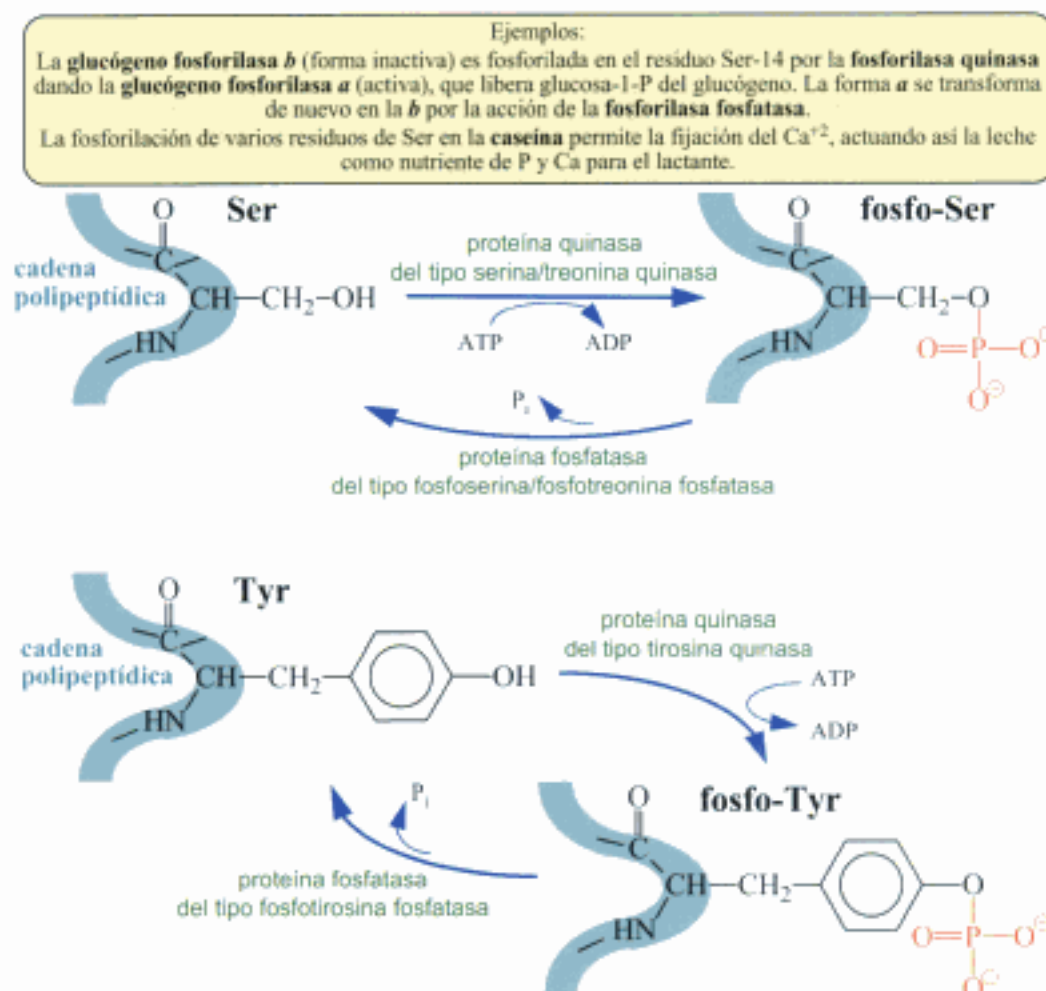
b) Carboxilación

En algunas proteínas se añade un grupo carboxilo a la cadena lateral de un aminoácido, por ejemplo al $-\text{CH}_2-$ de Asp (dando β -carboxiaspartato) y Glu (γ -carboxiglutamato, en los factores de coagulación y en proteínas estructurales del hueso).



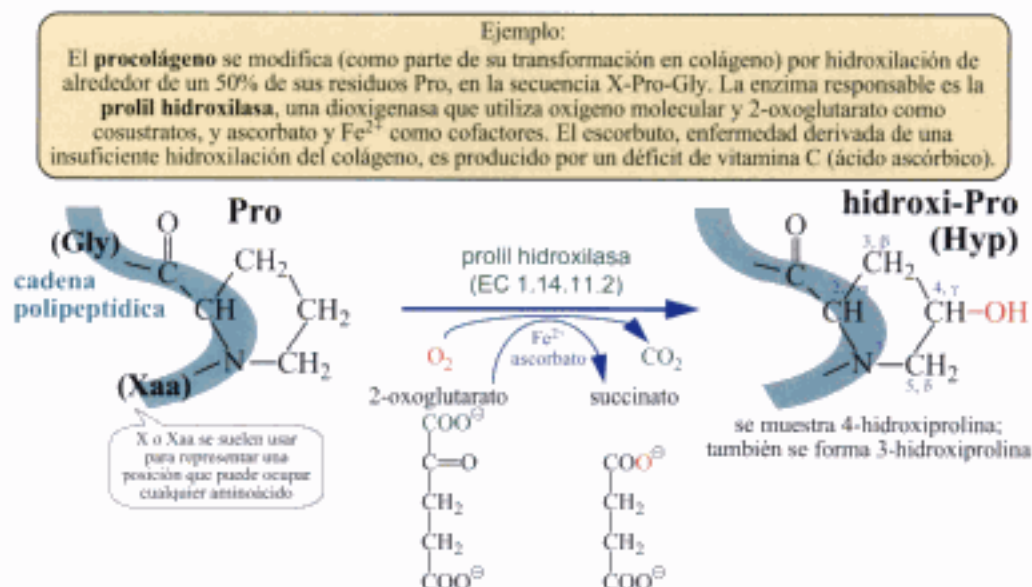
c) Fosforilación

Es quizá la modificación más frecuente, actuando casi siempre de forma reversible (**fosforilación-desfosforilación**) como un mecanismo regulador de la actividad de la proteína (regulación por **modificación covalente**). Afecta al grupo $-\text{OH}$ de Ser, Thr y Tyr, estando catalizada por proteína quinasas, con hidrólisis del ATP. Produce un incremento notable en la carga negativa de la proteína. La reacción opuesta, de desfosforilación, la catalizan proteína fosfatasa. En ambos casos, se trata de una modificación estrictamente postraduccional, que tiene lugar en el citosol después de que se ha completado la síntesis y el plegamiento de la proteína.



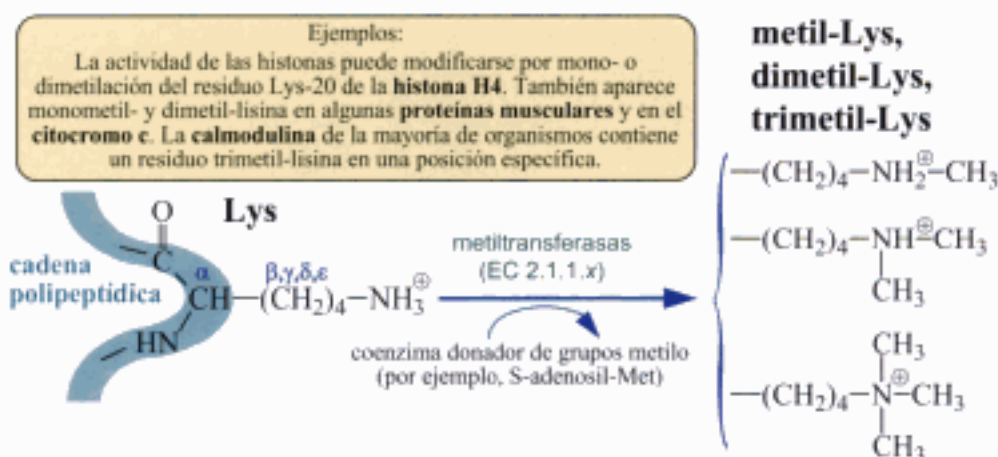
d) Hidroxilación

Varias hidroxilasas, presentes en el retículo endoplásmico, catalizan la incorporación de grupos -OH a algunas proteínas; por ejemplo, en Pro y Lys del colágeno.



e) Metilación

Consiste en la incorporación de metilo al grupo α -amino de la cadena lateral de Lys, o bien al grupo γ -carboxilo de Glu.



24.3.1.2 Incorporación de glúcidos: glicosilación

Se llama glicosilación a la unión covalente de cadenas de **oligosacáridos** (o **glicanos**) a la cadena lateral de algunos aminoácidos en las proteínas (actualmente se prefiere el prefijo **glico-** al antiguo **gluco-**, que queda reservado para lo específico de glucosa). Esta modificación es muy frecuente en proteínas de membrana y en aquellas que la célula secreta al exterior (entre las que se incluyen las proteínas plasmáticas), y menos común en proteínas intracelulares. Por otra parte, es característica de eucariotas y virus, pero inexistente en procariotas.

a) Características de las glicoproteínas

Las glicoproteínas en su conjunto no tienen una función común definida por la glicosilación. Sin embargo, la adición de azúcares tiene un papel esencial en la estructura y en la función de cada una de ellas. Por ejemplo, la glicosilación:

- Contribuye a establecer la conformación de la proteína.
- Aumenta su estabilidad (por ejemplo frente al calentamiento) y su resistencia a la digestión por proteasas. Como consecuencia, aumenta la vida media en plasma (disminuye el aclaramiento).

- Hace las proteínas más hidrofílicas, aumentando su solubilidad y, por tanto, facilitando su interacción con las células del entorno.
- Aporta estructuras individualizadas que median la interacción con receptores específicos y son importantes determinantes antigénicos.

Son innumerables las proteínas que están glicosiladas (más que las no glicosiladas), así como sus funciones, y no se discutirán en este texto. Como meros ejemplos, que no representan el conjunto de funciones, cabe citar su papel como base de los distintos tipos de **grupos sanguíneos**, entre ellos el grupo ABO (pág. 369), cuyos antígenos A, B y O son oligosacáridos constituyentes de glicoproteínas y glicolípidos presentes en la superficie de los eritrocitos y de otras células, y en glicoproteínas de secreción. Por otro lado, un método rutinario de vigilancia de la diabetes es la medida de **hemoglobina glicosilada**, cuya concentración depende de la de glucosa en sangre (la hemoglobina une covalentemente la glucosa que atraviesa libremente la membrana eritrocitaria). Igualmente, existen abundantes ejemplos de patologías provocadas por la alteración del patrón de glicosilación.

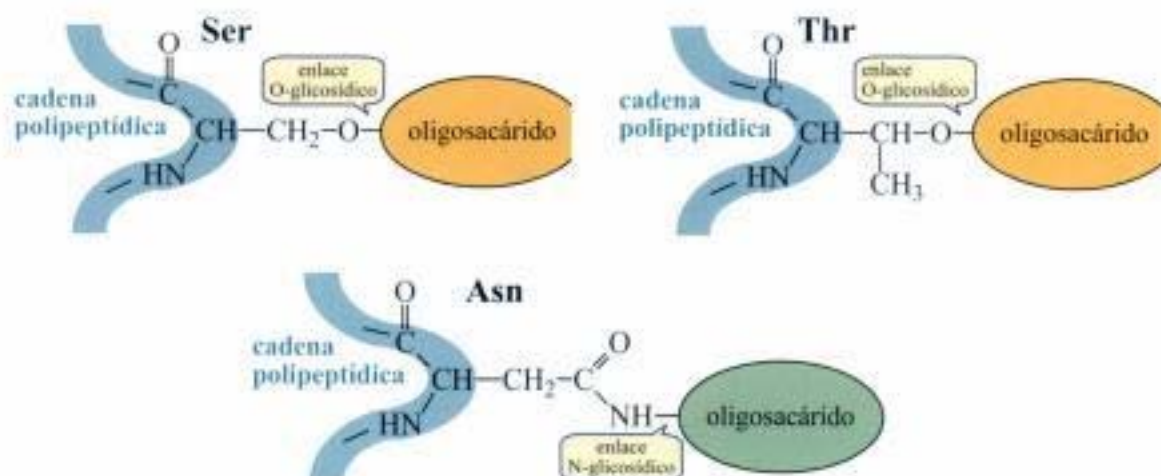
La unión de los oligosacáridos, que tiene lugar tanto de forma cotraduccional como postraduccional, da lugar a una gran variedad de glicoproteínas en las que la parte glucídica representa una proporción también variable de la masa total. La clasificación de los compuestos que poseen glúcidos y polipéptidos como parte de su molécula no es sencilla, debido a la gran diversidad existente. En cualquier caso, existen diferencias tanto en cuanto a la proporción de las partes glucídica y peptídica como al tipo de azúcares que participan y la forma como se unen entre sí. A partir de aquí, la descripción se limitará a lo relativo a las glicoproteínas propiamente dichas.

Moléculas que combinan polipéptido y glúcido

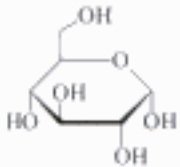
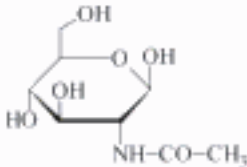
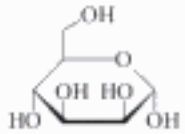
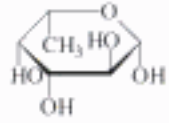
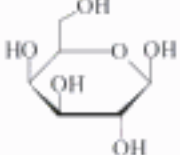
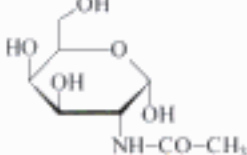
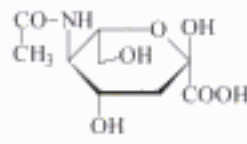
Glicoproteínas típicas o propiamente dichas	Glicanos de tamaño medio (oligosacáridos), ramificados e irregulares La parte glucídica supone del 2 al 85% de la masa total.
Proteoglicanos	La parte glucídica son glicosaminoglicanos , polisacáridos largos, lineales y regulares (repetición de un disacárido). Contienen una proporción significativa de aminoazúcares, sustituyentes sulfato y ácidos urónicos (monosacáridos con su grupo CH_2OH terminal oxidado a COOH). La parte glucídica supone un 85-90% de la masa total.
Colágenos	Un solo tipo de glicano. La parte glucídica supone menos del 1% de la masa total. Presentan propiedades mecánicas (elasticidad, resistencia, viscosidad, lubricación en articulaciones, etc.).
Peptidoglicanos	Característicos de la pared celular bacteriana. No hay proteína propiamente dicha. La unión es por enlace amida (en lugar de glicosídico).

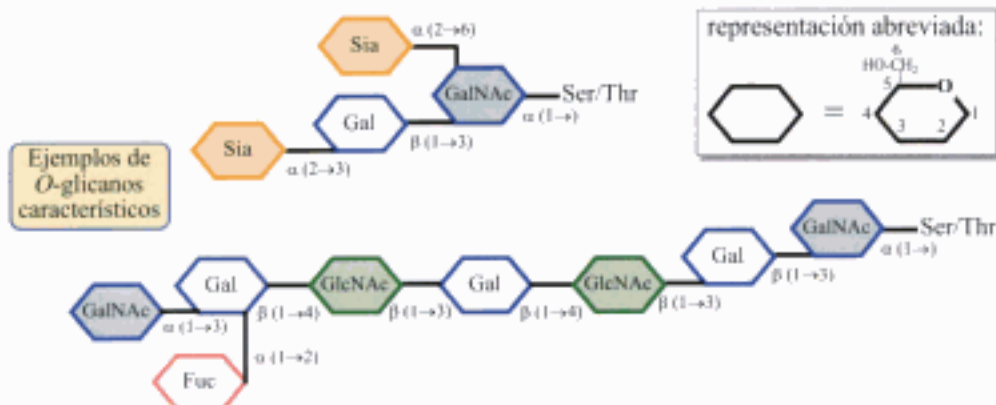
b) Tipos de glicosilación

Las cadenas glucídicas de las glicoproteínas se clasifican en dos grandes grupos, en función de cuál sea el enlace entre el oligosacárido y el polipéptido. Esta clasificación puede aplicarse en cierta medida a las propias glicoproteínas, pero algunas presentan simultáneamente azúcares unidos mediante ambos tipos de enlace, en puntos distintos de su molécula.

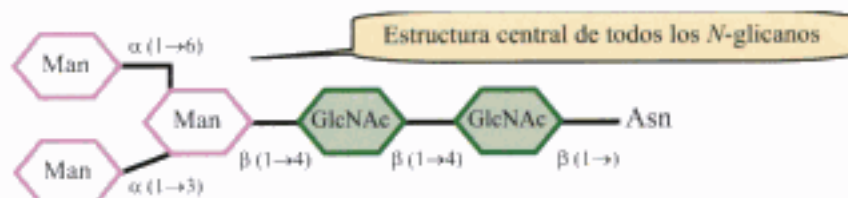


- **Oligosacáridos unidos por oxígeno (O-glicoproteínas):** el azúcar se une al átomo de oxígeno de cadenas laterales de Ser o Thr. En general, los O-glicanos, oligosacáridos O-unidos o unidos por O son estructuras sencillas, aunque variadas. Los más frecuentes tienen como base un residuo de N-acetilgalactosamina unido a serina o treonina, pero existen uniones xilosa-serina, galactosa-hidroxilisina, arabinosa-hidroxiprolina, N-acetilglucosamina-serina/treonina e incluso galactosa-cisteína (con un enlace S-glicosídico).

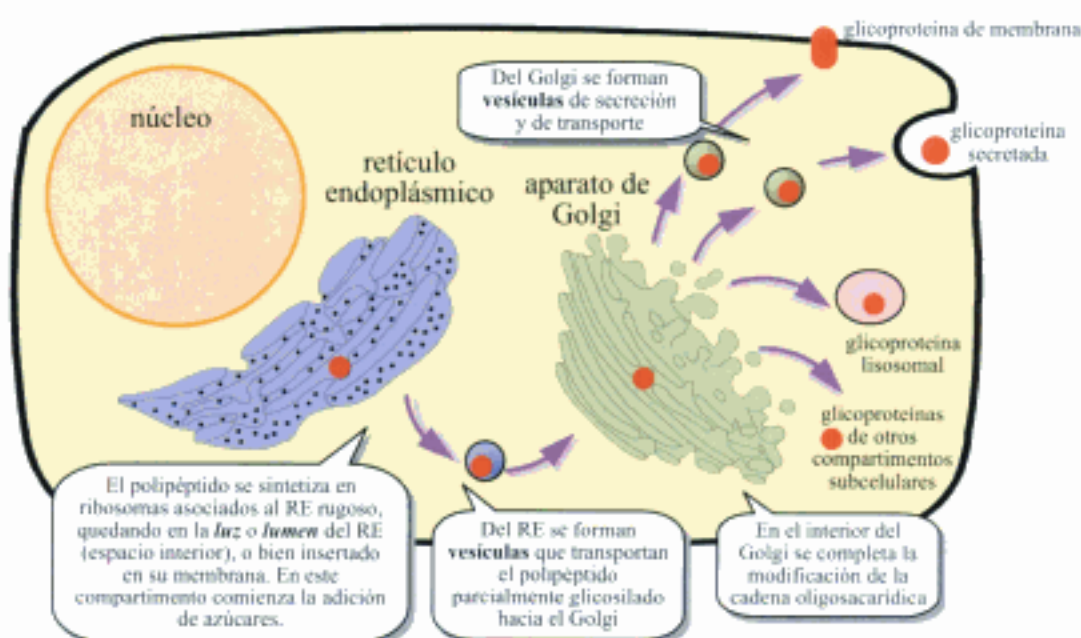
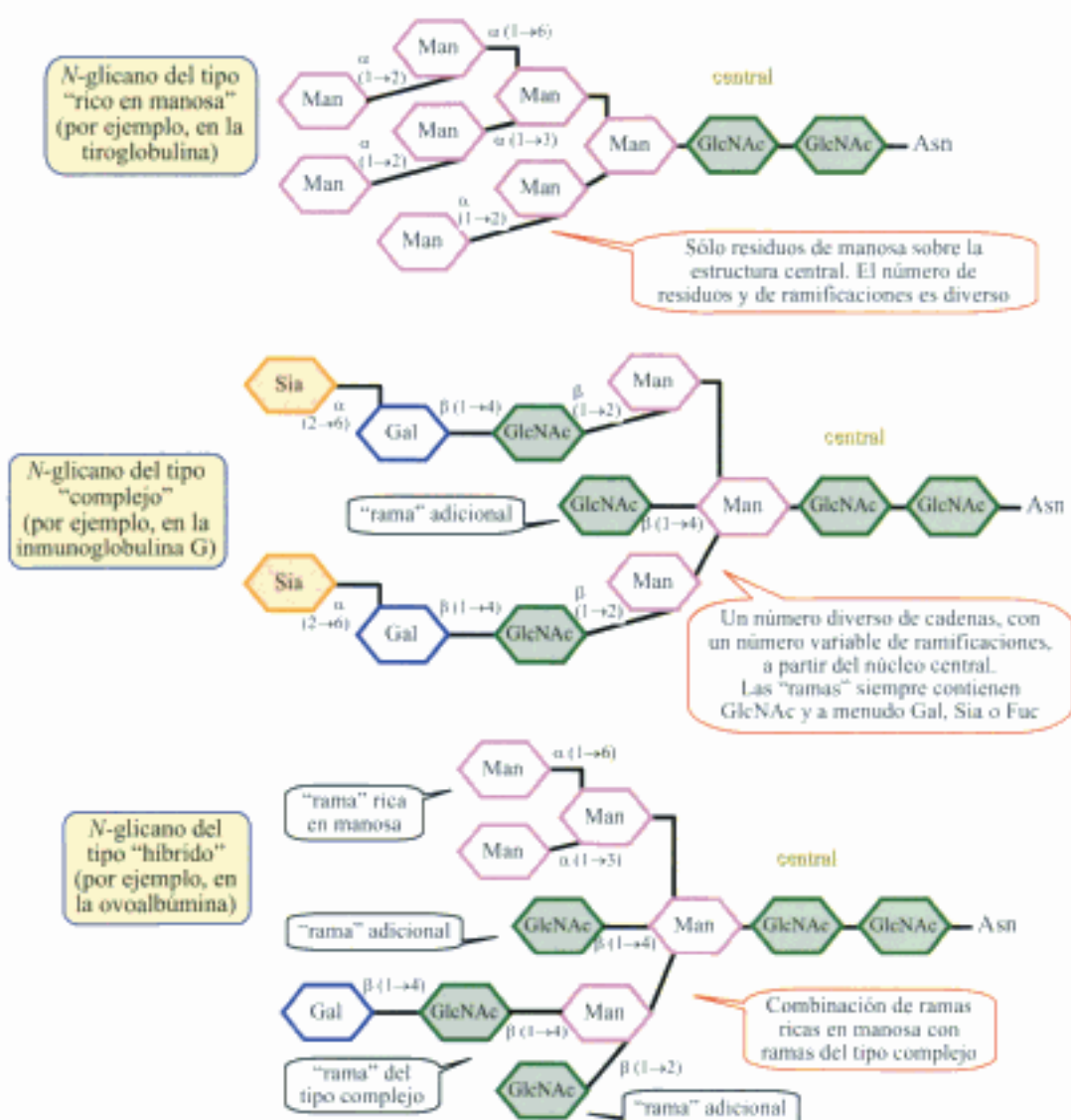
monosacárido:	α -D-Glucosa	N-acetil- β -D-glucosamina	α -D-Manosa (o β)	α -L-Fucosa (6-desoxi-L-Gal)
abreviatura:	Glc	GlcNAc	Man	Fuc
				
monosacárido:	β -D-Galactosa	N-acetil- α -D-galactosamina	Ácido Siálico (N-acetilneuramínico)	
abreviatura:	Gal	GalNAc	Sia (NeuNAc)	
				



- **Oligosacáridos unidos por nitrógeno (N-glicoproteínas):** el azúcar se une al átomo de nitrógeno amídico de la cadena lateral de Asn. Ésta debe formar parte de la secuencia Asn-X-Ser-Y, donde la posición de Ser puede estar ocupada también por Thr o Cys, y X e Y son cualquier aminoácido excepto Pro. Las estructuras de los N-glicanos, oligosacáridos N-unidos o unidos por N son mucho más complejas y diversas. El tipo de monosacáridos constituyentes es generalmente diferente del de los O-glicanos.



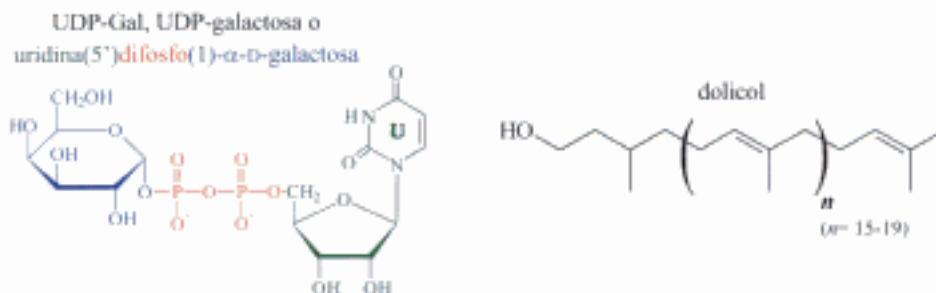
Alrededor de una estructura central común se unen monosacáridos muy diversos (ácido siálico, galactosa, N-acetilglucosamina ...), formando cadenas con distintas disposiciones. De acuerdo con éstas, los N-glicanos se clasifican en 3 grupos: ricos en manosa, complejos e híbridos.



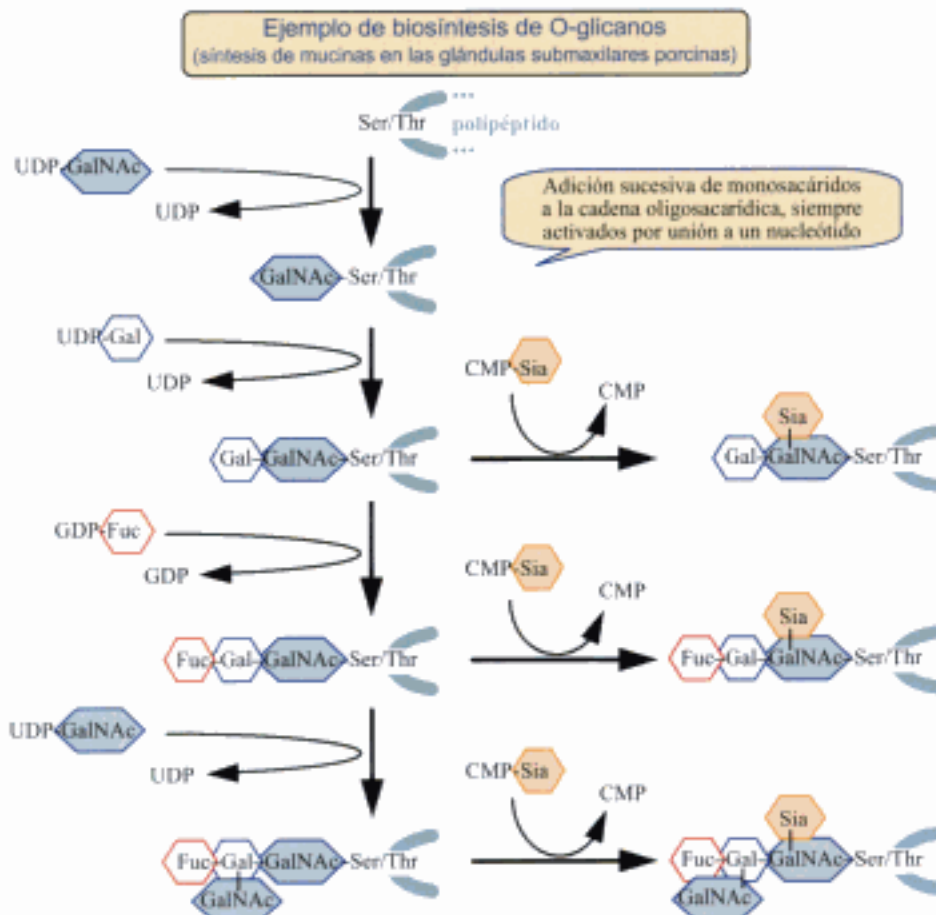
c) Biosíntesis de las glicoproteínas

La incorporación de las cadenas oligosacáridicas al polipéptido es uno de los principales procesos metabólicos que tienen lugar en el lumen del retículo endoplásmico (RE) y del aparato de Golgi. Las proteínas, por tanto, se mantienen aisladas del citoplasma durante su glicosilación, que tiene lugar mientras la proteína va viajando a lo largo de la ruta retículo endoplásmico → vesículas de transferencia → aparato de Golgi → vesículas de secreción o destino final de la glicoproteína (pág. 331).

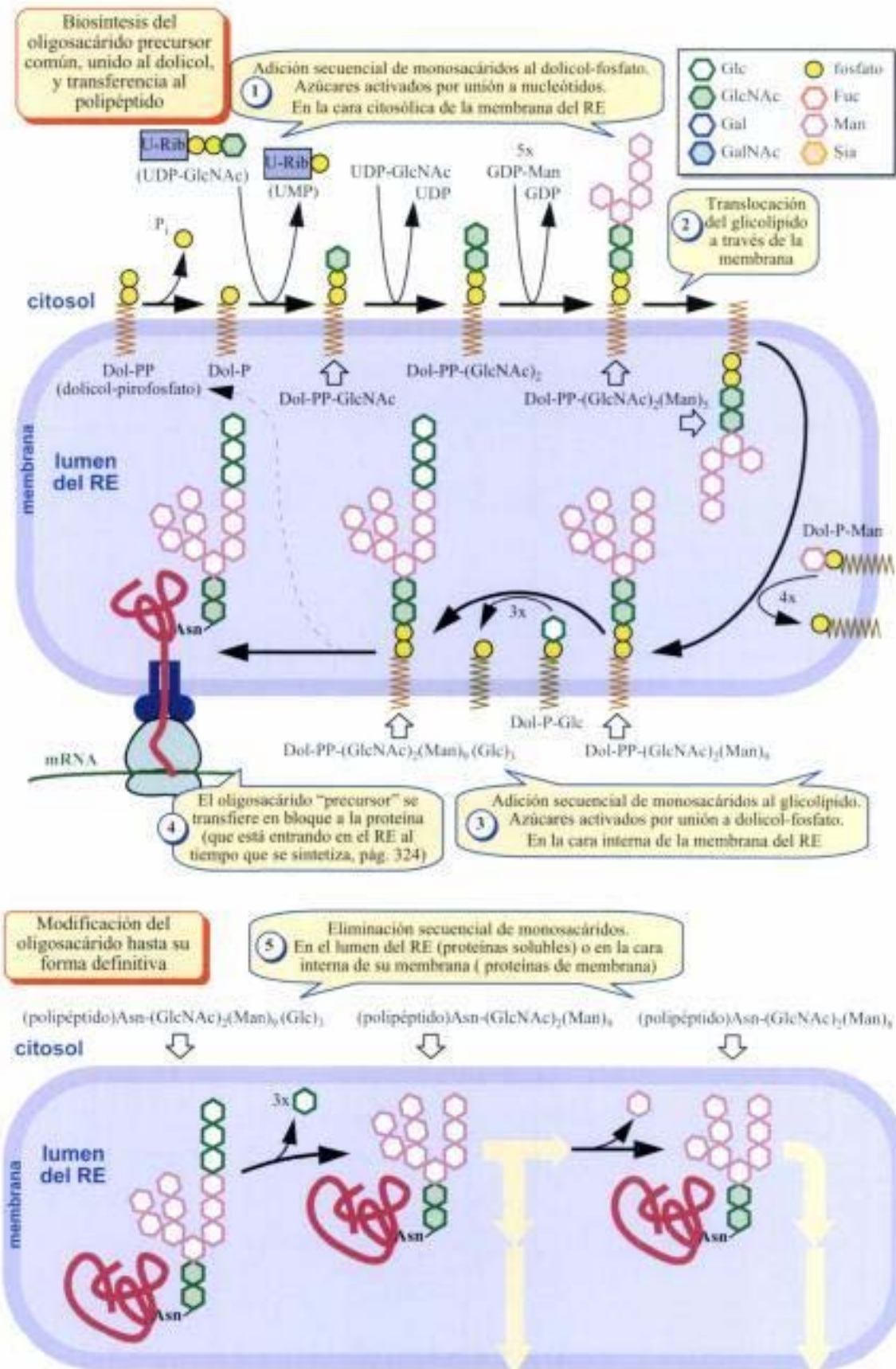
Los O-glicanos y N-glicanos se sintetizan por rutas metabólicas claramente diferentes, pero en todos los casos la unión de los monosacáridos está catalizada por **glicosiltransferasas**, específicas no sólo para reconocer el monosacárido que se incorpora y la cadena oligosacárida que actúa de aceptor, sino también para formar el enlace en posiciones muy definidas del anillo de cada uno de los azúcares. Estas enzimas se encuentran en el interior del retículo endoplásmico y del aparato de Golgi, asociadas a su membrana. Para poderse emplear como sustratos, los monosacáridos deben estar activados por unión a nucleótidos; concretamente, como UDP-Glc, UDP-GlcNAc, UDP-Gal, UDP-GalNAc, GDP-Man, GDP-Fuc y CMP-Sia. En otros casos, el precursor es el azúcar unido a una molécula de dolicol (un lípido terpenoide).

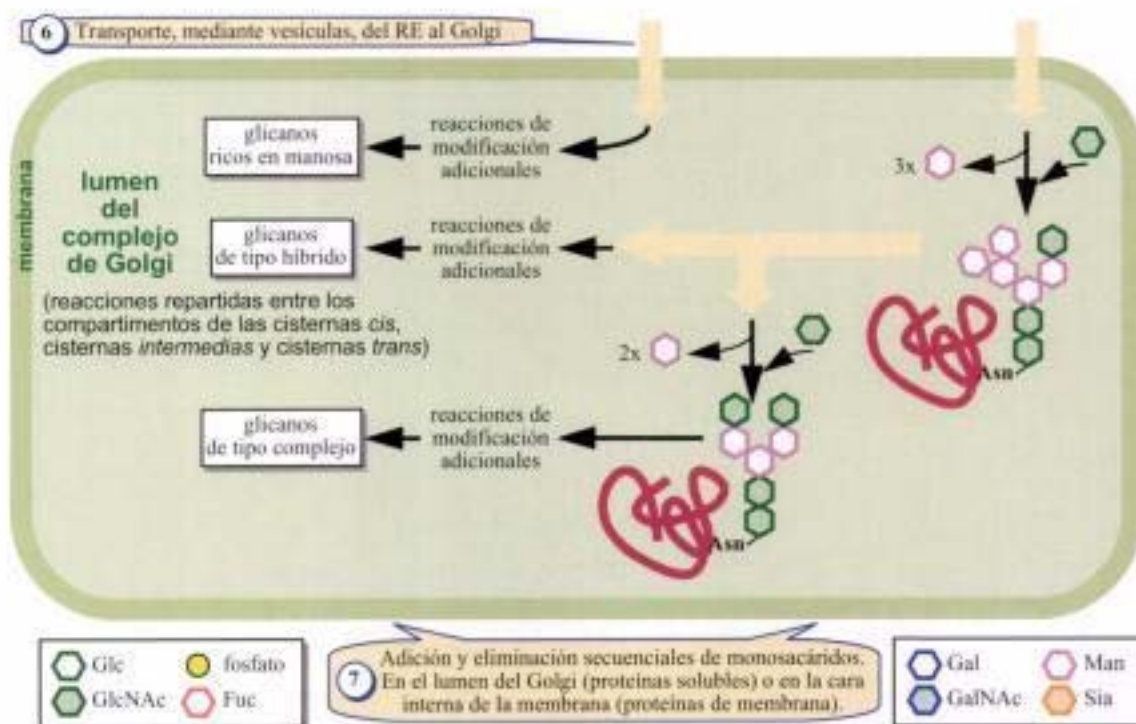


- La adición de los O-glicanos al residuo de Ser o Thr se realiza directamente, por unión sucesiva de los monosacáridos, catalizada por las respectivas glicosiltransferasas. Todas las etapas tienen lugar en el aparato de Golgi, de forma postraduccional (la síntesis de la cadena polipeptídica se termina previamente, en el RE).



- La síntesis de *N*-glicoproteínas es un proceso más complejo, que comienza por la síntesis de un oligosacárido precursor unido al dolicol, seguida de la transferencia *en bloque* de ese glicano al residuo de asparragina del polipéptido y finalmente de nuevas reacciones de adición de monosacáridos y también de eliminación de algunos de los previamente incorporados. El proceso se desarrolla secuencialmente entre el retículo endoplásmico (cotraduccionalmente) y el aparato o complejo de Golgi (postraduccionalmente).



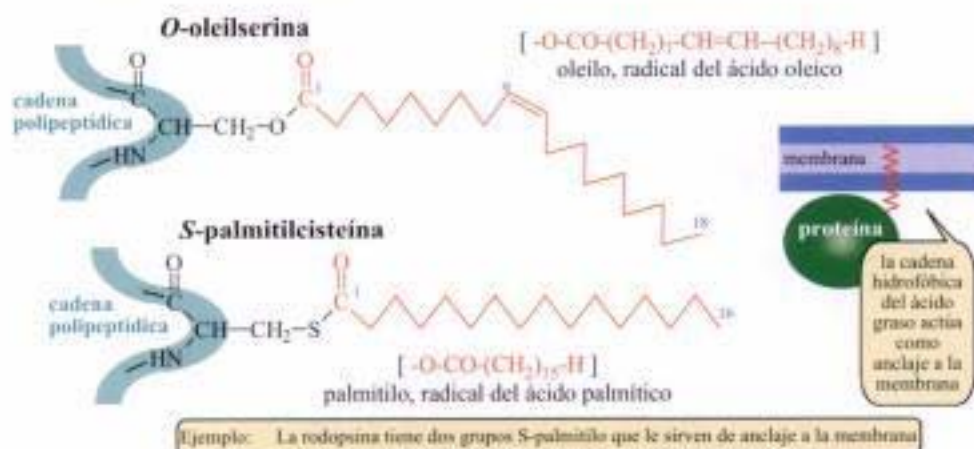


24.3.1.3 Modificación con lípidos

a) Acilación (ácidos grasos)

Aunque menos conocida que la glicosilación, la acilación es una modificación igualmente importante en algunas proteínas. Puede afectar a las cadenas laterales o a los extremos terminales del polipéptido; en general aumenta la hidrofobicidad de la proteína, y en concreto el lípido supone a menudo un punto de anclaje a la membrana (comúnmente en la cara interna). Generalmente esta reacción tiene lugar en proteínas citosólicas solubles, sintetizadas en ribosomas libres.

- Acilación de cadenas laterales de Ser y Thr (enlace éster con el grupo OH) o de Cys (enlace tioéster con el SH), con ácidos grasos de 14C (miristato), 16C (palmitato) y 18C (estearato y oleato).

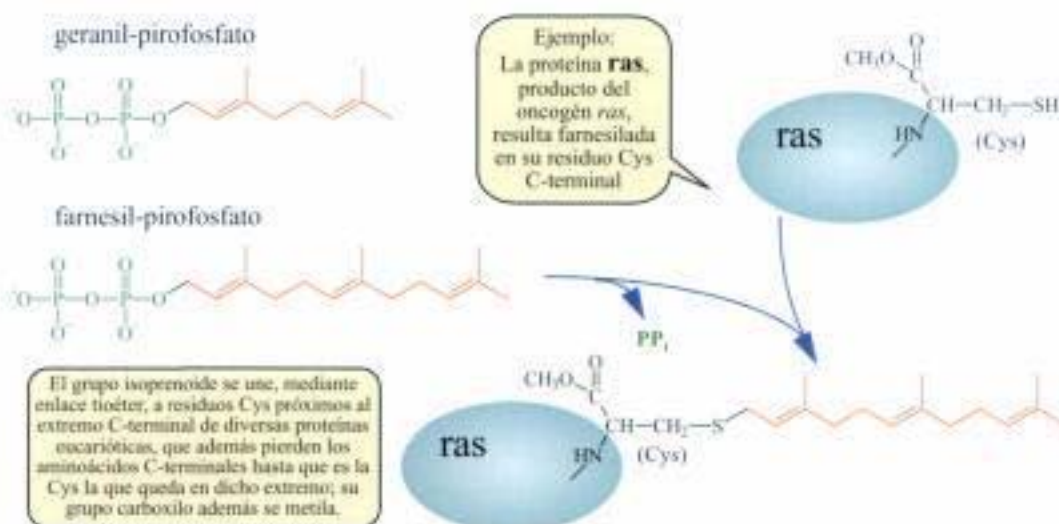


- **Acilación de los extremos.** El grupo amino terminal puede formar enlace amida con miristato (14C) o palmitato (16C), lo que tiene lugar cotraduccionalmente (antes de completar la síntesis del polipéptido). La reacción parte de miristil-coenzima A, catalizada por una miristil-CoA : proteína N-miristil transferasa:



b) Prenilación (terpenos)

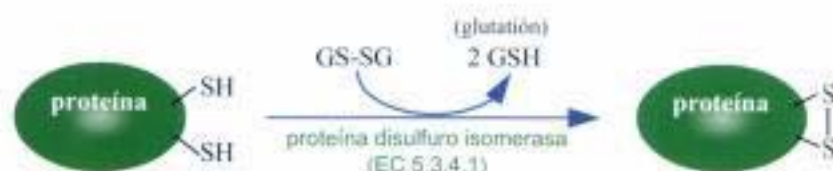
Principalmente se unen tres tipos de radicales terpenoides (formados por unidades de isopreno; de ahí el nombre "prenilación"): geranilo (10C), farnesilo (15C) y geranilgeranilo (20C); los dos primeros son metabolitos intermediarios de la ruta biosintética del colesterol. Al igual que la acilación, sufren esta modificación proteínas sintetizadas en ribosomas citosólicos.



Entre las proteínas modificadas con farnesilo se encuentran la proteína **Ras**, proteínas G y proteínas de la matriz nuclear. La oncoproteína **Ras** utiliza el farnesilo para anclarse a la superficie interna de la membrana plasmática y poder activar las señales intracelulares que desembocan en el fenotipo canceroso. El bloqueo de la isoprenilación conlleva la pérdida de la actividad transformante (carcinogénica), por lo que se han desarrollado fármacos anticancerosos basados en la inhibición de la farnesil-transferasa.

24.3.1.4 Formación de puentes disulfuro

Muchas proteínas forman entrecruzamientos covalentes mediante puentes disulfuro entre restos Cys de la misma o de distinta cadena (intra- o intercatenarios, respectivamente). Estos enlaces, que se dan con menor frecuencia en proteínas intracelulares que en las que la célula secreta, participan en el plegamiento correcto (pág. 338) y protegen la conformación nativa de la proteína frente a la desnaturalización en un ambiente extracelular que puede ser más agresivo. La formación de enlaces disulfuro (una reacción redox) viene ayudada por la enzima **proteína-disulfuro isomerasa** y por moléculas pequeñas con grupos SH (tiol), la principal de las cuales es el glutatión (γ -glutamilcisteinilglicina).



24.3.1.5 Otras modificaciones

- Para ejercer su actividad muchas proteínas procarióticas y eucarióticas emplean **grupos prostéticos**, algunos de ellos unidos covalentemente una vez que el polipéptido ha abandonado el ribosoma. Entre estos grupos pueden citarse la biotina en la acetil-CoA carboxilasa o el grupo hemo en mioglobina, hemoglobina y citocromo c.
- La actividad de algunas enzimas se regula mediante la adición de **mononucleótidos**. Por ejemplo, la glutamina sintetasa (del sistema fijador de nitrógeno en procariotas) en su forma inactiva se encuentra adenilada en un residuo Tyr. El grado de adenilación es controlado por una proteína reguladora.
- La **ADP-ribosilación** es otra reacción de modificación que actúa de forma reversible con carácter regulador de la función de proteínas citoplásmicas y nucleares. Transferasas específicas actúan sobre His, Arg, Asn, Lys o Glu, empleando como cosustrato el NAD⁺. La misma actividad es ejercida, perturbando la regulación fisiológica, por las toxinas diftérica, cólerica y pertúsica.
- En algunas proteínas se produce la **sulfatación** de Tyr con 3'-fosfoadenosina-5'-fosfosulfato; tiene lugar en el complejo de Golgi.

24.3.2 Maduración por escisión de las proteínas: Procesamiento proteolítico

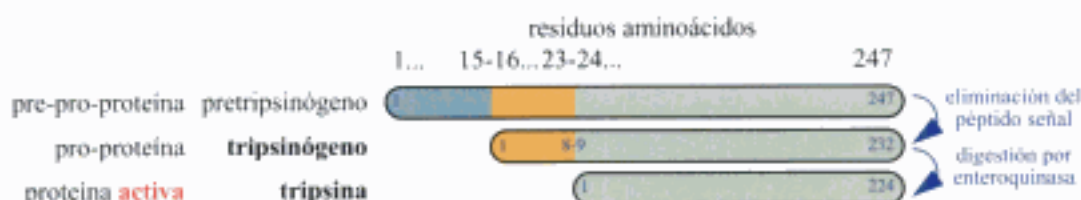
Este segundo tipo de maduración proteica consiste en la ruptura específica del polipéptido (**proteólisis**) en polipéptidos más pequeños, cada uno con distinta actividad, o en la eliminación de una porción sin función (generalmente pequeña) para dar un solo polipéptido activo. En cualquier caso, se trata de un proceso de **activación**, de precursores inactivos a proteínas activas. No debe confundirse esta rotura con la responsable de la degradación proteica (pág. 341).

Además de los procesos de activación por proteólisis, que se describen a continuación, se deben considerar dentro de este apartado la eliminación de residuos amino y carboxi-terminales, esencialmente la liberación enzimática de la metionina N-terminal (previa desformilación en proteínas procarióticas), y la eliminación de la llamada **secuencia señal** (responsable de la dirección o destino de las proteínas maduras, pág. 324).

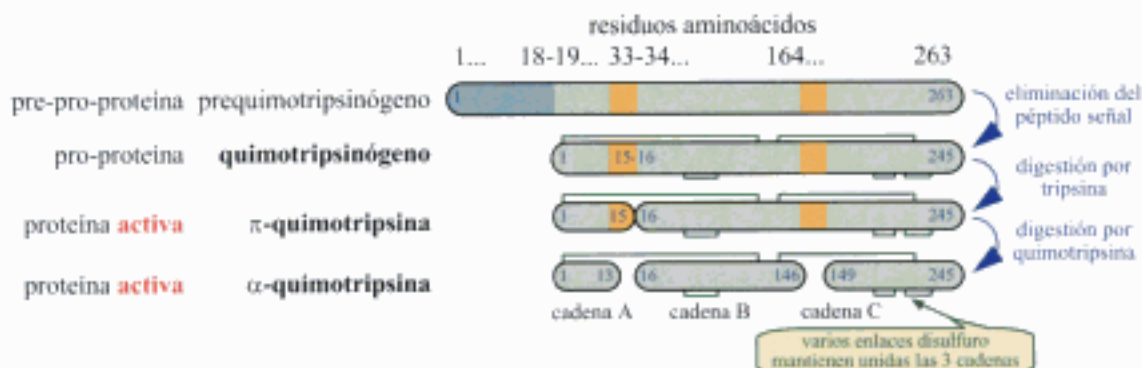
24.3.2.1 Activación de proenzimas proteolíticas

Las proteasas digestivas **tripsina** y **quimotripsina** se sintetizan en forma inactiva, como precursores o pro-proteínas, denominadas **zimógenos** (en realidad, recién sintetizadas son **pre-pro**-proteínas, pues poseen un **péptido señal** que las marca como proteínas de secreción y luego es eliminado). De ese modo se protegen las células donde se sintetizan y se reserva la actuación para el tubo digestivo.

La **tripsina**, producida en el páncreas, se activa por eliminación de su octapéptido N-terminal, por la acción de la **enteroquinasa** producida por las células de la pared del intestino delgado:



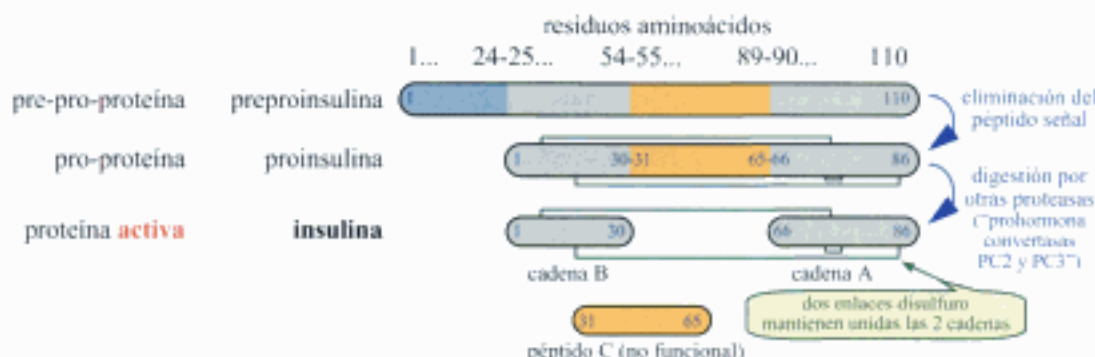
La tripsina actúa a su vez activando otras proenzimas o zimógenos, entre ellos el **quimotripsinógeno**, precursor de otra enzima digestiva pancreática, la quimotripsina; esta activación es algo más compleja:



24.3.2.2 Activación de precursores de hormonas peptídicas

Como ejemplo característico, veamos la síntesis de la **insulina**. Además de demostrar la maduración por escisión proteolítica, constituye un ejemplo típico de modificación química por formación de enlaces disulfuro (pág. 335).

1. La insulina se sintetiza en las células B (β) de los islotes de Langerhans del páncreas bajo la forma de **preproinsulina**, un precursor polipeptídico inactivo. Se evita así la posibilidad de que la insulina activa pudiera actuar sobre dichas células. La presencia en su extremo amino de una secuencia señal determina, como en toda proteína destinada a la secreción, el transporte de esta molécula mediante vesículas hacia el exterior celular.



2. En el retículo endoplásmico se elimina la secuencia señal dando la **proinsulina**, el verdadero precursor aún inactivo. Éste se pliega en una conformación tridimensional específica, estable, que facilita la formación de tres enlaces disulfuro.
3. La proinsulina se transporta hasta en el aparato de Golgi, donde comienza su conversión por proteólisis en **insulina activa**. Las moléculas se acumulan en vesículas denominadas **gránulos de secreción**, donde se completa su proteólisis y se almacenan hasta que la célula recibe las señales que determinan la secreción de la insulina al medio extracelular.

24.4 PLEGAMIENTO DE PROTEÍNAS

Uno de los problemas importantes en biología molecular es comprender cómo la información contenida en la secuencia de aminoácidos de una proteína da lugar a su estructura tridimensional distintiva, responsable de su función o actividad biológica. Este **plegamiento** de la cadena polipeptídica, que tiene lugar en unos segundos, se realiza simultáneamente al tráfico y a una parte de la maduración del polipéptido, por lo que maduración y plegamiento están íntimamente ligados.

Recientemente se ha descubierto que los defectos en las rutas de plegamiento y tráfico de las proteínas son la causa molecular de varias enfermedades.

- La **fibrosis quística** se origina por una mutación en el gen de la proteína transportadora de cloruro, que suprime uno de sus aminoácidos (pág. 351). Este cambio impide la separación del complejo proteína-carabina (pág. 338), lo que a su vez impide que la proteína salga del RE y continúe su maduración y tráfico hacia la membrana citoplasmática.
- La **enfermedad de Alzheimer**, que afecta al 10% de las personas mayores de 65 años, se origina en parte por la acumulación de agregados de un péptido derivado de la proteína β -amiloide; esos agregados posiblemente están causados por un defecto en el plegamiento.
- Una forma hereditaria de **enfisema** (enfermedad pulmonar degenerativa) se produce por mutaciones en el gen de la α_1 -antitripsina (pág. 376) que conducen a la agregación de la proteína, impidiendo su secreción; la deficiencia de la proteína en plasma produce susceptibilidad al desarrollo de la enfermedad.
- Por último, la conocida enfermedad de las vacas locas, **scrapie** en ovejas y **enfermedad de Creutzfeld-Jacob** en humanos son causadas por una proteína con plegamiento anómalo, la proteína priónica PrP. Aunque la descripción de este problema supera las posibilidades de este libro, baste decir que una molécula de proteína plegada anormalmente induce el mismo plegamiento en otras moléculas de la misma proteína, conduciendo así por un lado al desarrollo de la enfermedad y por otro a su transmisión a otros individuos.

24.4.1 Mecanismo del plegamiento de las proteínas

Algunos polipéptidos, especialmente los más pequeños, pueden adoptar su conformación tridimensional nativa de forma espontánea, pero en la mayor parte de los casos el plegamiento correcto requiere la ayuda de proteínas especiales (distintas de las que se están plegando); la traducción del término inglés *chaperones* ha llevado a llamarlas **chaperonas**, **chaperoninas**, **celadoras** o **carabinas**; emplearemos este último nombre, más fiel al significado original (persona de edad acompañante de una joven, que evita "interacciones incorrectas"). Al parecer no sólo están implicadas en el plegamiento, sino también en la translocación a través de membranas subcelulares y en la degradación de las proteínas.

Además de las carabinas, en las células contribuyen al plegamiento dos tipos de isomerasas que, al catalizar ruptura y reconstrucción de enlaces covalentes, aceleran el plegamiento sin afectar a su resultado final. Entre ellas cabe citar la **proteína-disulfuro isomerasa**, que intercambia enlaces disulfuro entre dos moléculas (pág. 335) y otras enzimas que catalizan la interconversión de las conformaciones *cis* y *trans* de la prolina formando enlace peptídico, contribuyendo ambas a que la cadena polipeptídica adopte distintas conformaciones en la búsqueda de la de mínima energía.

El conocimiento de los mecanismos e interacciones que estabilizan la estructura tridimensional surge de experimentos pioneros de desnaturalización y renaturalización proteica (con la enzima ribonucleasa, de pequeño tamaño y un solo dominio) y posteriormente de la aplicación de métodos biofísicos de análisis y tecnología del DNA para producir polipéptidos alterados. Dada la variedad de proteínas y la complejidad de los mecanismos implicados, no se conocen aún las reglas que rigen el plegamiento *in vivo*: aún no somos capaces de predecir la estructura tridimensional de una proteína a partir de su secuencia.

Sin embargo, está claro que es la propia secuencia de aminoácidos la que determina por sí misma la estructura tridimensional activa. La estructura primaria o secuencia lineal resultante de la asociación por enlaces peptídicos de los aminoácidos (determinada a su vez por la secuencia lineal de bases del DNA a través del mRNA) define la disposición de la cadena polipeptídica para formar puentes de hidrógeno, la capacidad de sus residuos hidrofóbicos para acomodarse en el interior de la molécula y los requerimientos estéricos que definen las posiciones más idóneas de la conformación nativa. Este punto tiene implicaciones de índole práctica, biotecnológica, porque cualquier modificación de la secuencia altera la estructura final de la proteína.

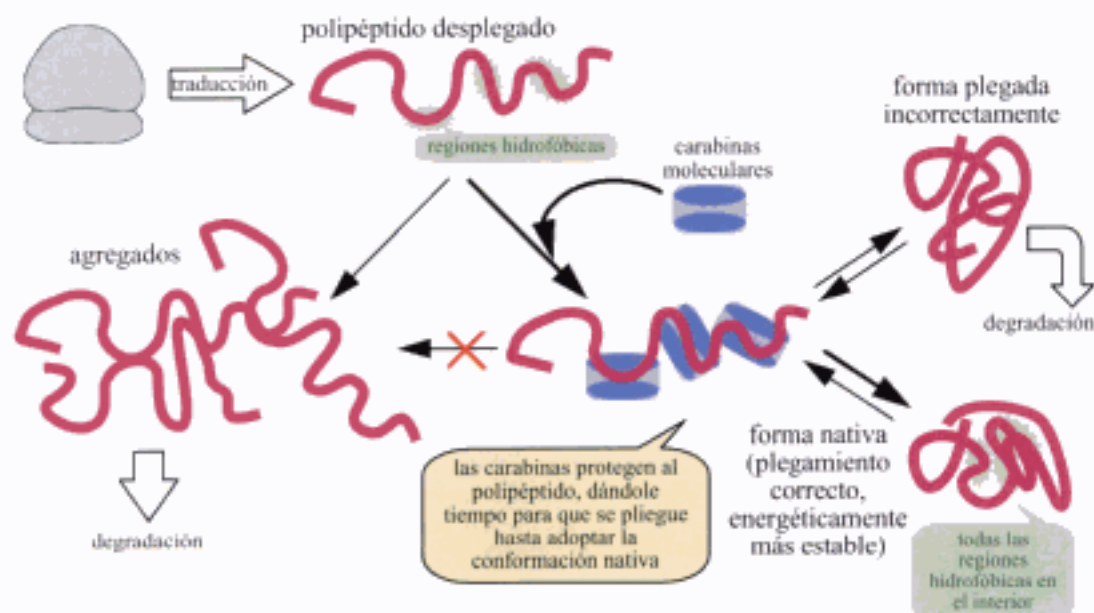
Las etapas del proceso de plegamiento pueden analizarse, de forma aproximada, siguiendo los niveles de organización de las proteínas globulares. En general, el plegamiento viene determinado por consideraciones energéticas correspondientes a las estructuras secundaria y terciaria, aquellas que permiten el máximo número de puentes de hidrógeno e interacciones de van der Waals, iónicas e hidrofóbicas, es decir, la conformación de menor energía (termodinámicamente, la más estable).

Muchas proteínas se pliegan a través de estados globulares intermedios, conocidos como de **"glóbulo fundido"** (*molten globule*), estructuras compactas aunque más abiertas y flexibles que la conformación nativa, con una estructura secundaria considerable, más o menos similar a la de aquella, pero que carecen aún de su estructura terciaria definida. Las cadenas laterales de los aminoácidos en el glóbulo fundido no han establecido aún sus interacciones mutuas, presentan libertad de movimiento e interacción con el disolvente, especialmente las hidrofóbicas. Este primer plegamiento sería como el "esqueleto" sobre el que se elaboraría la estructura tridimensional definitiva, algo así como un "pliegue en partes" de la estructura final.

El principio básico del plegamiento, en su conjunto, es la presencia de residuos hidrofóbicos de la cadena en un ambiente acuoso altamente polar: el "efecto hidrofóbico", el mismo que determina la asociación de los lípidos en las membranas celulares, es la principal fuente de estabilidad de la conformación plegada de una proteína.

24.4.2 Proteínas implicadas en el plegamiento: carabinas moleculares

Los polipéptidos en estado de glóbulo fundido, los plegados parcial o incorrectamente y especialmente aquellos que aún están saliendo del ribosoma (cadenas nacientes) exponen en su superficie residuos hidrofóbicos que en la proteína nativa estarán siempre dispuestos en el interior de la molécula ("escondidos" o "enterrados" por la envoltura de residuos hidrofílicos que forma la superficie de la proteína). Estas regiones hidrofóbicas, para evitar la interacción con el medio acuoso polar, tienden a asociarse entre sí, conduciendo a la formación de agregados de varias moléculas (que en



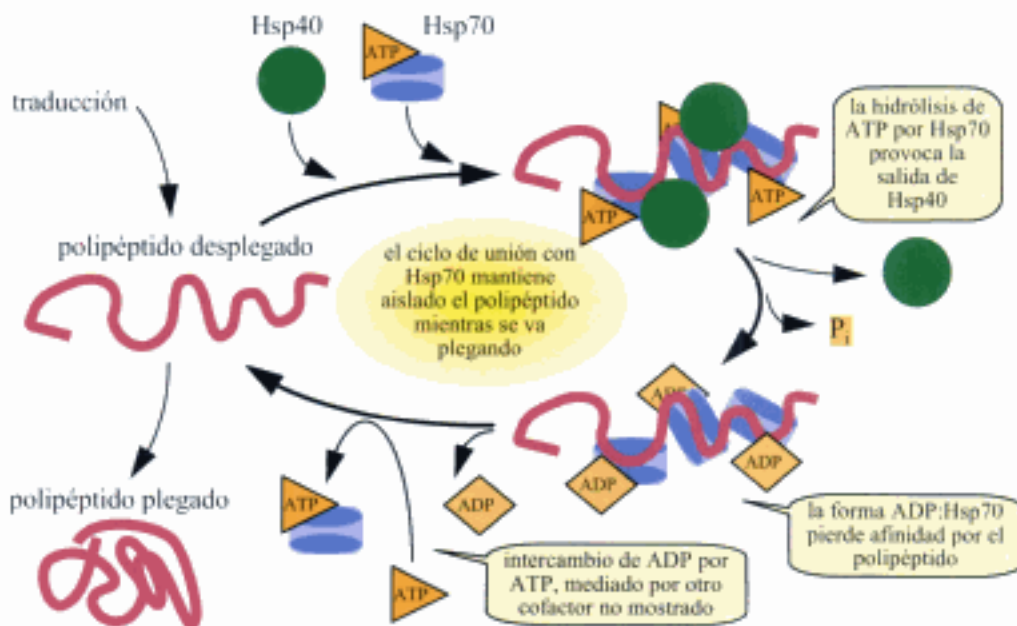
algunos casos constituyen los llamados "cuerpos de inclusión"). El plegamiento correcto, por tanto, requiere una protección frente a esta posibilidad, hasta que la proteína tenga la ocasión de encontrar su conformación nativa, la más estable; éste es uno de los papeles de las **carabinas moleculares**.

24.4.2.1 Tipos de carabinas

Las primeras proteínas identificadas como accesorias, auxiliares o catalizadoras del plegamiento se llamaron **proteínas de choque térmico** (**Hsp**, por *heat shock protein*), porque su síntesis se estimula al someter células en cultivo a una elevación temporal de la temperatura. Esto es así porque en esas condiciones aumenta la cantidad de moléculas parcialmente desplegadas, que es necesario replegar. Actualmente, el grupo de carabinas incluye otras muchas proteínas, no relacionadas entre sí, importantes para la viabilidad celular a todas las temperaturas y bajo todas las circunstancias. De acuerdo con su modo de actuación, se pueden distinguir dos grandes grupos de carabinas moleculares:

a) Carabinas del tipo Hsp70

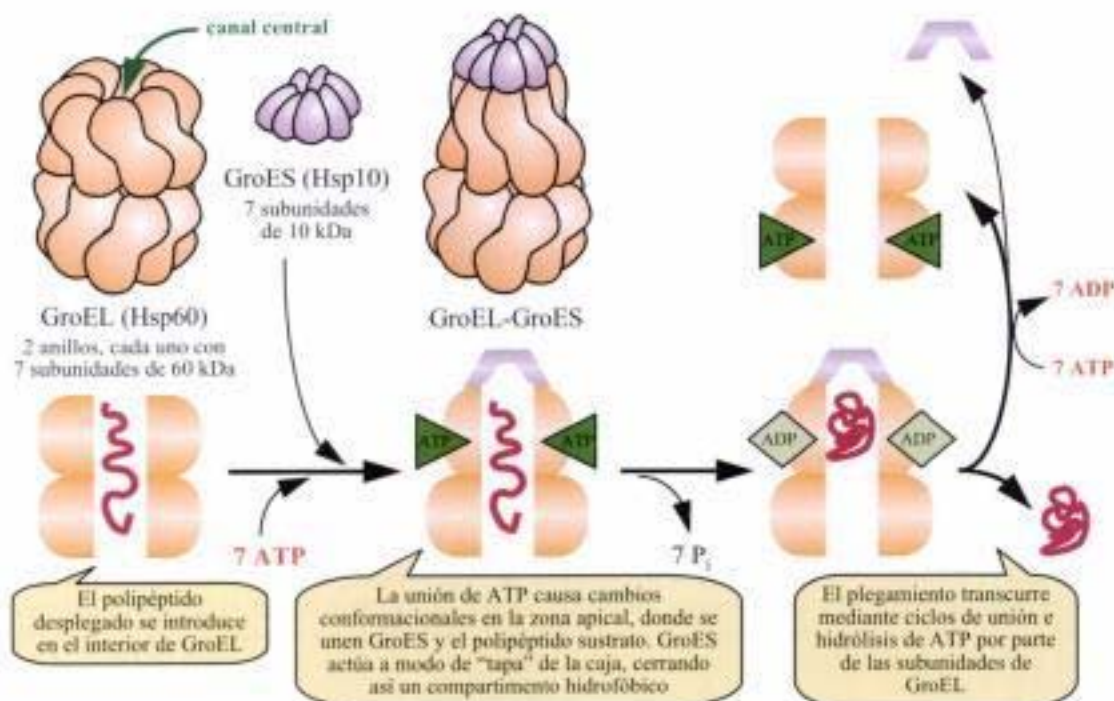
Esta familia (cuyo nombre procede de su masa molecular de 70 kDa) comprende proteínas que se asocian a segmentos peptídicos cortos, en conformación extendida o desplegada y ricos en aminoácidos hidrofóbicos. Por ello, reconocen las cadenas polipeptídicas nacientes según están saliendo del ribosoma. Su acción principal es evitar que se agreguen por interacciones hidrofóbicas intermoleculares (recordando la metáfora de la señorita). Existen carabinas de esta familia en procariotas, eucariotas, mitocondrias, cloroplastos y en el interior del retículo endoplásmico. Su unión y separación, reversibles, con la proteína desplegada dependen de la hidrólisis de ATP por parte de una actividad ATPasa propia de la Hsp70. Estos procesos están a menudo modulados por otro tipo de proteínas, la familia **Hsp40**, con capacidad de activadores de la Hsp70 (co-carabinas) y también de carabinas moleculares por sí mismas.



Las carabinas Hsp70 estabilizan, por tanto, las cadenas polipeptídicas nacientes, protegiéndolas de la agregación, y mediante los ciclos controlados de unión y separación proporcionan un entorno "seguro" para que la proteína pueda plegarse adecuadamente. Sin embargo, no contribuyen a definir el plegamiento, que sólo depende de la secuencia del polipéptido.

b) Carabinas del tipo Hsp60

Estas proteínas forman una estructura multimérica muy característica, en forma de caja cilíndrica, en cuyo interior se alojan los polipéptidos y donde experimentan su plegamiento, aislados del entorno y de nuevo evitando la interacción con otras moléculas desplegadas o parcialmente plegadas. El ejemplo mejor conocido es la proteína GroEL de *E. coli*, cuya estructura se ha determinado por cristalografía de rayos X a alta resolución. Con la colaboración de GroES, proteína de la familia **Hsp10**, dirige el plegamiento del polipéptido. Si éste no se pliega completamente en un ciclo, puede volver a entrar en la cavidad para continuar plegándose, hasta alcanzar la estructura nativa.



Se han encontrado homólogos del sistema GroEL-GroES en mitocondrias y cloroplastos, y proteínas menos estrechamente relacionadas en el citosol eucariótico.

24.4.2.2 Función de las carabinas

Las carabinas moleculares desempeñan un papel esencial en varias circunstancias:

- En el citosol celular, dada la gran densidad de orgánulos y macromoléculas, la variedad y elevada concentración de proteínas allí presentes en fase de plegamiento aumenta la probabilidad de interacciones hidrofóbicas y de agregación de las proteínas, que, como se ha comentado, es evitada por las carabinas.
- En la mitocondria y el cloroplasto intervienen en el desplegamiento necesario para el tráfico de proteínas de origen citosólico a través de las membranas externa e interna del orgánulo, así como en el plegamiento en su interior, para adquirir su funcionalidad.
- En el lumen del RE interactúan con las proteínas plegadas incorrectamente (o las proteínas dimericas o multiméricas no ensambladas de forma adecuada), con lo que éstas quedan retenidas en el RE hasta que tiene lugar el plegamiento correcto. Sólo entonces (al igual que las proteínas bien plegadas desde el principio) se transportan en vesículas hacia el complejo de Golgi. En caso contrario, las proteínas no plegadas son degradadas.

24.5 DEGRADACIÓN DE LAS PROTEÍNAS

En paralelo con el desempeño de su función, las proteínas se degradan constantemente en la célula. Este proceso continuo, unido a la biosíntesis, constituye el llamado **recambio proteico** (en inglés, *turnover*), que tiene lugar a lo largo de toda la vida celular y es necesario para mantener en todo momento un nivel adecuado de proteínas (estado estacionario), indispensable para la actividad celular, para impedir la acumulación de proteínas anormales o no necesarias y para facilitar el reciclado de los aminoácidos. De forma genérica, se puede decir que en una célula una proteína "sobrevive" en promedio unos dos días antes de ser degradada, lo que da idea de la gran actividad de los procesos de biosíntesis y de degradación.

Existen proteínas que se sintetizan y degradan (se recambian) de forma muy rápida respecto a la vida media de la célula donde se encuentran. Entre ellas se encuentran las proteínas defectuosas, por ejemplo por haber incorporado aminoácidos incorrectos en la síntesis, las que no desempeñan correctamente su función; a veces un simple cambio químico convierte a una proteína en blanco de una acción proteolítica. Por el contrario, otras proteínas, las necesarias en todo momento, son muy estables y por ello no requieren una síntesis continua a elevada velocidad. Así ocurre, por ejemplo, con la hemoglobina, que puede durar todo el tiempo de vida media del eritrocito (120 días en humanos).

En eucariotas existen dos sistemas principales de degradación de las proteínas, uno mediante los lisosomas y el otro en el citosol.

24.5.1 Degradación lisosómica

El primer sistema de degradación descrito se realiza en los lisosomas, liberados inicialmente por gemación a partir del complejo de Golgi (lisosomas primarios) y que después reciben enzimas hidrolíticas mediante vesículas procedentes también del Golgi. Estos orgánulos delimitados por una membrana sencilla acumulan más de 50 enzimas hidrolíticas de todo tipo, capaces de actuar sobre proteínas, ácidos nucleicos, lípidos, enzimas glicolíticas, etc., con una gran actividad al pH ácido característico del lisosoma (alrededor de 5). Las enzimas lisosomales que hidrolizan proteínas (proteasas) se denominan **catepsinas**.

Las enzimas hidrolíticas concretas presentes en los lisosomas dependen de cada tipo de célula; como ejemplo, las células pancreáticas contienen lisosomas con enzimas capaces de intervenir en las siguientes rutas de degradación:

- Para la exocitosis, o transporte de su contenido enzimático al exterior de la célula (secreción).
- Para la fagocitosis o formación de lisosomas fágicos para digerir orgánulos.
- Para destruir vacuolas fagocíticas con nutrientes de la superficie celular.
- En ocasiones, para romper intracelularmente los lisosomas y liberar sus enzimas, conduciendo a la autólisis de la célula (una de las formas de muerte celular programada, que interviene por ejemplo en el desarrollo embrionario).

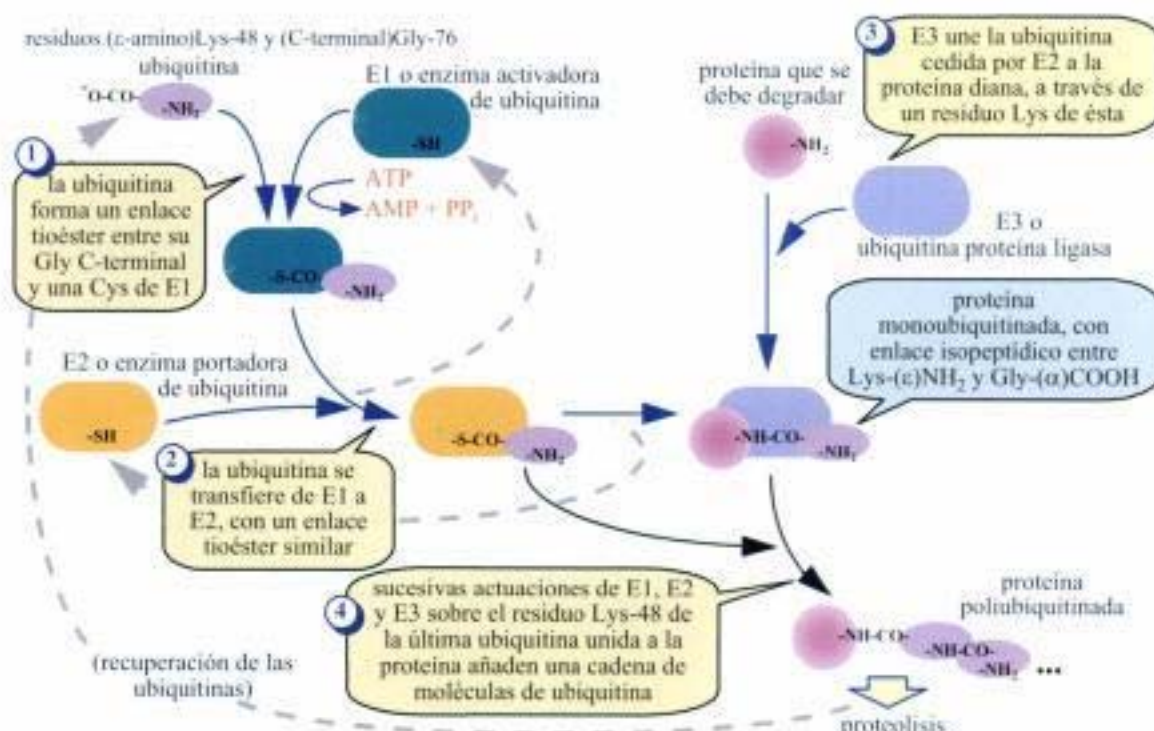
24.5.2 Degradación citosólica

La mayoría de las proteínas presentes en forma libre en el citosol se pueden degradar por esta ruta. El proceso es muy selectivo, está bajo un estricto control y es la ruta de degradación que requiere mayor gasto de ATP. Al igual que en el tráfico de proteínas, intervienen como mediadores en la degradación señales moleculares especiales, concretamente aminoácidos N-terminales y la proteína **ubiquitina**.

24.5.2.1 Ubiquitina

La ubiquitina es una proteína pequeña, de 76 aminoácidos, llamada así por encontrarse en todas las células eucariotas (es ubicua). Es posiblemente una de las proteínas más conservadas entre especies (las de levadura y humanos sólo se diferencian en tres aminoácidos). De carácter globular, rígida y muy estable, interviene como marcador selectivo de la degradación de otras proteínas, al unirse covalentemente a ellas, de forma que las proteínas ubiquitinadas quedan dispuestas para ser digeridas por los sistemas proteolíticos.

La **ubiquitinación**, ubiquitinilación o unión de la ubiquitina a cualquier proteína destinada a la destrucción se realiza en un proceso de tres reacciones, catalizadas por 3 enzimas diferentes:



El enlace amida de la ubiquitina con la proteína (llamado isopeptídico porque no tiene lugar con el amino α, sino con el ε de la cadena lateral de Lys) constituye la señal para la destrucción, el marcaje para la proteólisis (se le ha llamado "el beso de la muerte"), especialmente cuando hay multiubiquitinación (hasta 20 moléculas de ubiquitina unidas en cadena a una sola de proteína).

24.5.2.2 Señales desencadenantes de la ubiquitinación

a) Residuo amino-terminal

La identidad del aminoácido N-terminal parece determinar que una proteína resulte marcada con ubiquitina con mayor o menor facilidad. Recuérdese que en eucariotas este residuo resulta de la eliminación de Met y de cualquier otro proceso de maduración de dicho extremo; las señales amino-terminales parecen haberse conservado durante miles de millones de años en la evolución.

aminoácido en posición N-terminal	vida media de la proteína
grupo I: básicos: Arg, Lys, His	Prolongada
grupo II: hidrofóbicos voluminosos: Leu, Ile, Trp, Phe, Tyr	Prolongada
grupo III: neutros: Gly, Ala, Val, Ser, Thr, Cys, Met, Pro	Corta

b) Secuencias PEST

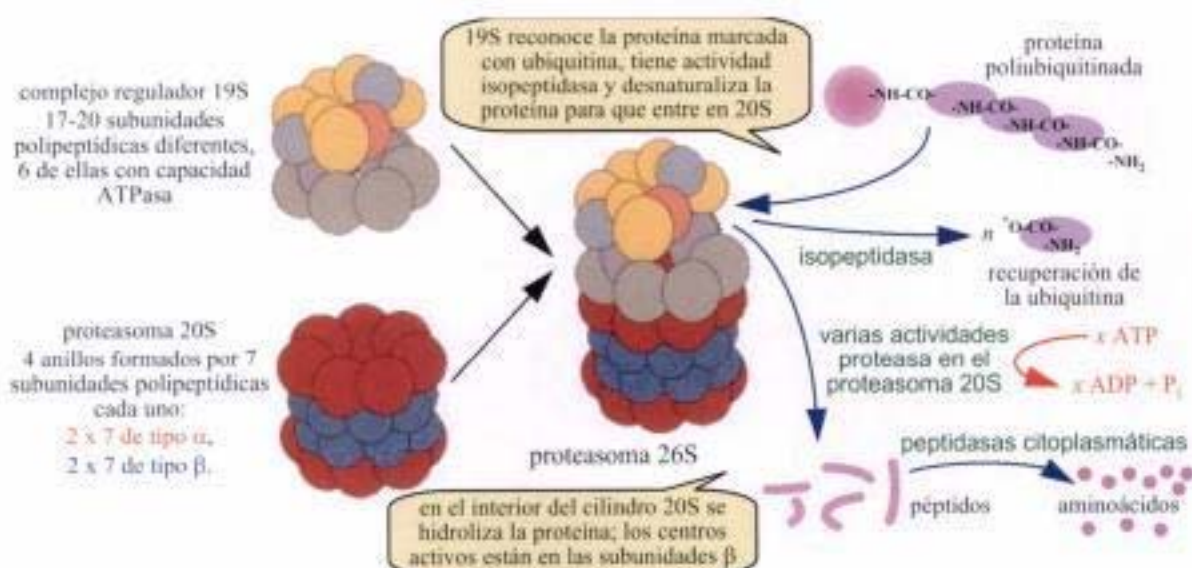
Por otra parte, también se asocia una vida media corta con la presencia en la proteína de regiones de 12 a 60 aminoácidos, ricas en Pro, Glu, Ser y Thr (P, E, S, T, con la abreviatura de una letra). Estas regiones actuarían a modo de esquema de reconocimiento para los sistemas enzimáticos que degradan las proteínas de vida corta.

24.5.2.3 Proteasas responsables de la degradación

La proteólisis de las proteínas marcadas con ubiquitina puede ser responsabilidad de distintos sistemas de proteasas. Se han descrito **calpainas**, proteasas que requieren calcio para su actuación, pero el sistema más importante es probablemente el del proteasoma.

Se ha denominado **proteasoma** a un gran complejo proteico (2000 kDa) que aparece en muchas células eucarióticas, formado por distintas subunidades proteolíticas. El conjunto, **proteasoma 26S** o **complejo proteinasa multicatalítico**, contiene dos componentes asociados entre sí:

- Un componente central o **proteasoma 20S**, formado por 28 subunidades en una disposición cilíndrica, que es el núcleo catalítico responsable de la proteólisis secuencial, con liberación de fragmentos peptídicos y de las unidades de ubiquitina intactas para su reutilización. (En cierto modo, recuerda a la estructura de las carabinas Hsp60.)
- Uno o dos componentes laterales, **complejo 19S**, con actividad ATPasa y capacidad de reconocer específicamente los conjugados proteína-poliubiquitina.



Tema 25

Bases moleculares de la mutación y la reparación del DNA

25.1	Concepto de mutación	344
25.2	Clasificación de las mutaciones	344
25.3	Tipo de célula que sufre la mutación	344
25.3.1	Mutación en células de la línea germinal	344
25.3.2	Mutación en células somáticas	345
25.4	Mutaciones a pequeña escala o puntuales	346
25.4.1	Cambios en la secuencia del DNA	347
25.4.1.1	Sustitución	347
a)	Transiciones	347
b)	Transversiones	347
25.4.1.2	Delección, pérdida o eliminación	348
25.4.1.3	Inserción	348
25.4.2	Efecto sobre la secuencia de la proteína sintetizada	348
25.4.2.1	Mutaciones silenciosas	348
25.4.2.2	Mutaciones no silenciosas	349
a)	Mutaciones que alteran el sentido	349
b)	Mutaciones que cambian el marco de lectura	350
c)	Mutaciones que no cambian el marco	351
d)	Mutaciones con terminación prematura de la proteína	351
e)	Mutaciones con terminación retrasada	352
25.5	Causas y mecanismos básicos de las mutaciones	352
25.5.1	Mutaciones endógenas	352
25.5.1.1	Incorporación de nucleótidos erróneos durante la replicación	352
25.5.1.2	Otras mutaciones espontáneas	353
a)	Sustituciones por desaminación oxidativa	353
b)	Pérdida de bases por inestabilidad química del enlace N-glicosídico	354
c)	Modificación de las bases por mutágenos endógenos	354
25.5.2	Mutaciones inducidas o exógenas: mutagénesis	354
25.5.2.1	Lesiones o mutaciones inducidas por agentes químicos	354
25.5.2.2	Lesiones y mutaciones inducidas por agentes físicos	356
25.6	Reparación del DNA	356
25.6.1	Eliminación de agentes mutágenos	357
25.6.2	Reversión directa de la lesión	357
25.6.2.1	Reparación de fotodímeros	357
25.6.2.2	Reversión de bases modificadas con grupos alquilo	357
25.6.3	Reparación por escisión	358
25.6.3.1	Mecanismo general de reparación por escisión de nucleótidos	359
25.6.3.2	Mecanismos específicos de reparación por escisión de bases	360
25.6.4	Reparación de apareamientos incorrectos	362
25.7	Enfermedades humanas asociadas a la reparación	363

25.1 CONCEPTO DE MUTACIÓN

Clásicamente, se define la **mutación** como una alteración en la secuencia del DNA de un individuo que se transmite por herencia a sus descendientes. La forma inalterada del DNA, llamada tipo salvaje o silvestre (en inglés, *wild type*), se convierte así en otra forma, portadora de la mutación o tipo *mutante*. Las mutaciones se producen por errores en la replicación, por la alteración espontánea de nucleótidos o debido a la acción de agentes físicos o químicos (**mutágenos**), pero nunca como consecuencia de la recombinación meiótica entre cromosomas homólogos (pág. 106).

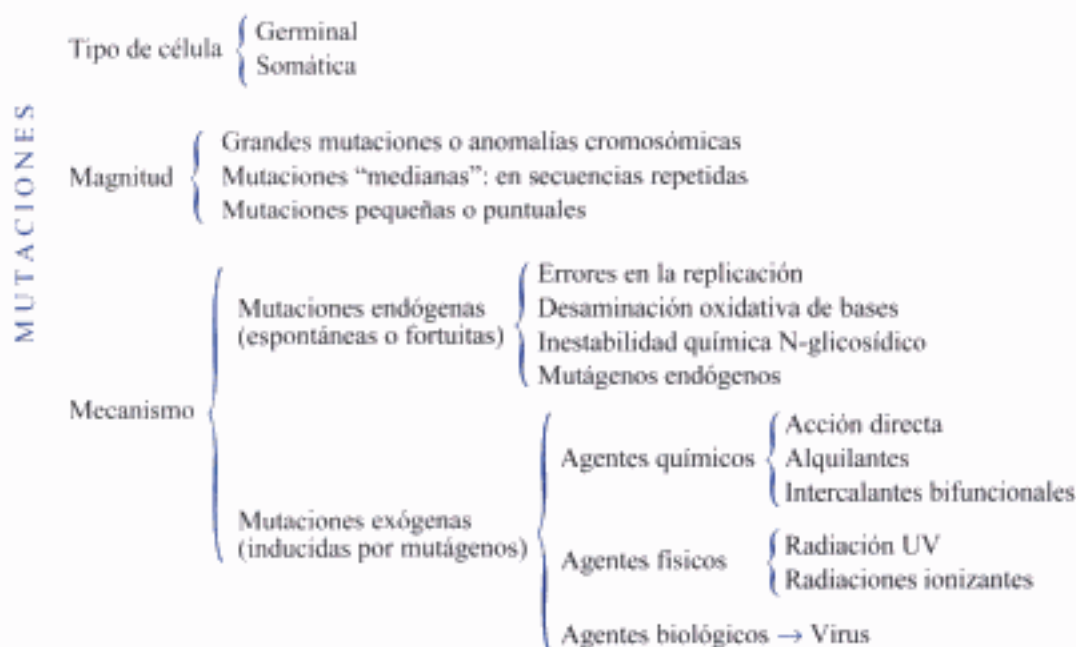
El carácter hereditario incluido en la definición supone que la mutación debe presentarse de modo estable en las células germinales. Ahora bien, ello no es estrictamente así. Aunque con frecuencia se asume que la mutación sólo corresponde a secuencias codificantes (genes), siendo las alteraciones en los productos génicos (generalmente proteínas) las responsables de las enfermedades hereditarias, el concepto de mutación es hoy día más amplio.

Las mutaciones tienen lugar en todo el genoma, incluyendo las secuencias codificantes o no, del genoma nuclear o mitocondrial, en células germinales (heredable) o somáticas (no heredable). Por tanto, la mutación es, junto con la recombinación meiótica (pág. 106), la principal fuente de variabilidad genética en todo tipo de organismos, y su estudio tiene interés básico y aplicado por su relación con procesos aparentemente tan alejados como la evolución de las especies, la identidad genética de un individuo, las alteraciones patológicas de su genoma responsables de enfermedades, el diagnóstico prenatal o el tratamiento de enfermedades somáticas (por ejemplo, el cáncer) o de células germinales.

Aunque muchas mutaciones se generan al azar, y existe la misma susceptibilidad de mutación en todas las regiones del genoma, las que ocurren sobre el DNA codificante (3% del total) tienen peores consecuencias. Dentro de esta **mutación génica** se debe distinguir entre la de la región del gen directamente responsable de la información (*región estructural*), que altera el RNA o la proteína sintetizados, y la que ocurre en las secuencias no codificantes del gen implicadas en el control de la expresión (*región reguladora*), que no altera la proteína sino su síntesis (aumento o descenso en concentración).

25.2 CLASIFICACIÓN DE LAS MUTACIONES

Dada la variedad de factores implicados, existen diversos planteamientos para clasificar las mutaciones. Se estudiarán bajo tres criterios: tipo de célula (germinal o somática), magnitud de la mutación (pequeña, mediana o grande) y mecanismos causantes de la mutación.



25.3 TIPO DE CÉLULA QUE SUFRE LA MUTACIÓN

25.3.1 Mutación en células de la línea germinal

Se origina en alguna de las divisiones mitóticas o meióticas de la gametogénesis (págs. 102-103), a partir de células progenitoras diploides de las gónadas. La mutación puede no influir en el fenotipo de la célula o individuo que la sufre, ni en la capacidad reproductora de éste, pero, puesto que el genoma se transmite a la descendencia, la mutación **se hereda**, transmitiéndose a **todas las células** del organismo hijo y generando, en su caso, la alteración o enfermedad

hereditaria. Obviamente, esa transmisión depende de que en la fecundación participe un gameto portador de la mutación. A su vez, las células germinales del hijo también la poseerán (al igual que todas sus células somáticas), por lo que se volverá a transmitir a la siguiente generación. En consecuencia, la mutación se comporta como una variación heredable, que estará sujeta a la selección natural; se puede decir que son la *materia prima* para los cambios evolutivos.

La importancia de estas mutaciones depende directamente de cuál sea la función afectada: si ésta es esencial, se originan organismos inviables (abortos tempranos, a veces no detectados), mientras que si la función no es esencial, pero sí importante, se origina una enfermedad congénita o hereditaria, o una subnormalidad. Sin embargo, no debe olvidarse que algunas mutaciones resultan beneficiosas y son la base de la evolución.

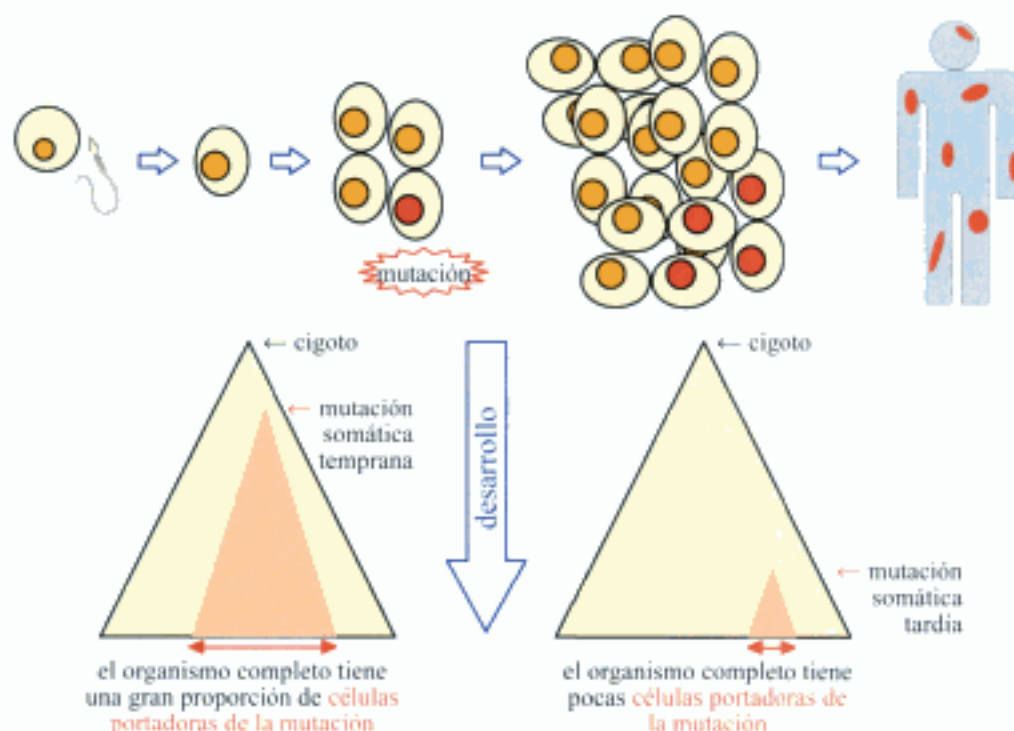
Las mutaciones en la línea germinal son habitualmente acontecimientos esporádicos, minoritarios y relativamente infrecuentes en la población humana. Su tasa es de unas 10^{-6} mutaciones por locus y división celular. A este respecto, se puede calcular el número de mutaciones que tienen lugar durante las divisiones mitóticas y meióticas de la espermatogénesis. En la formación de un espermatozoide humano aparecen alrededor de 100 mutaciones debidas a errores en la replicación, lo que, teniendo en cuenta el tamaño de su genoma haploide, supone una mutación por cada 33 Mb. La mayoría de ellas son recesivas o por el contrario tan letales que no son fenotípicamente evidentes, pues no alcanzan la concepción o el nacimiento. De ahí que la tasa de mutaciones de la línea germinal con consecuencias patológicas sea muy pequeña.

25.3.2 Mutación en células somáticas

A diferencia de las anteriores, las mutaciones somáticas (en células no reproductoras) sólo se transmiten a las células hijas (en la mitosis), y no a un organismo completo. Pueden afectar a cualquiera de los 46 cromosomas de cualquier célula somática, diploide, en cualquier órgano, tejido o célula aislada, durante el desarrollo precoz o final del organismo, o en su fase adulta. Estas mutaciones determinan las características subnormales o patológicas del individuo, pero los cambios **nunca son heredables**.

Éstas son mutaciones **mayoritarias** porque las células somáticas son las que más se dividen (unas 10^{13} divisiones en humanos desde el cigoto hasta el individuo adulto). La tasa de mutación, del orden de 10^{-6} por pb y división celular, es inferior a la de células germinales. Sin embargo, se incrementa como consecuencia de la exposición a agentes mutágenos (algunos de ellos cancerígenos), de la edad y de la alteración de los mecanismos de reparación del DNA, llegando a provocar miles de nuevas mutaciones durante la vida de un adulto, prácticamente en cada posición en el genoma y en cualquier parte del organismo.

El individuo portador de una mutación somática que se ha producido después del cigoto, pero antes de alcanzar su desarrollo completo, puede considerarse un *mosaico*, término genético que define la presencia de dos o más líneas celulares genéticamente distintas derivadas de un mismo cigoto. A través de la mitosis, las células somáticas del individuo adulto pertenecerán a dos líneas o clones, la *propia* o normal y la *anómala*, derivada de la célula que sufrió la mutación. Los individuos con una réplica fiel y completa de ambos tipos de células, muestran *mosaicismo somático*. Este fenómeno se puede observar, por ejemplo, en ciertos trastornos citogenéticos con alteración en el número de cromosomas (pág. 406), como la trisomía 21. Lógicamente, cuanto más temprano tiene lugar la mutación en el desarrollo del organismo, mayor será la abundancia de células de la línea anómala sobre las de la línea normal.



Las mutaciones que sean compatibles con la vida de la célula pueden, sin embargo, afectar a su crecimiento. El efecto de la mutación sobre la célula somática depende en parte de su compatibilidad con el ciclo de división celular, pues la célula afectada se divide activamente y conducirá a la proliferación de la mutación, pudiendo ocasionar problemas clínicos. Así ocurre, por ejemplo, en procesos de envejecimiento celular y transformación neoplásica. El **cáncer** se inicia por la mutación de genes que controlan la división celular (tema 30).

25.4 MUTACIONES A PEQUEÑA ESCALA O PUNTUALES

El estudio de las mutaciones en función de su magnitud, longitud o grado de la lesión permite clasificarlas en pequeña y gran escala (y, a veces, mediana). Las **anomalías a gran escala** ocurren por duplicaciones, pérdidas o reagrupamientos que alteran una región de millones de pares de bases y se reflejan en la estructura del cromosoma. Se llaman también **mutaciones cromosómicas** o **voluminosas**. Su estudio se reserva para el tema 29, como anomalías cromosómicas. En éste nos limitaremos a las mutaciones a pequeña escala.

Las **mutaciones a pequeña escala, sencillas, puntuales, mínimas o micromodificaciones**, implican normalmente a un solo nucleótido (y, como consecuencia, a su complementario en la otra hebra del DNA). Sin embargo, también se incluyen aquí aquellas otras lesiones a pequeña o mediana escala que afectan a varios pb. Nuestro estudio pretende cubrir, esencialmente, los tipos de mutación puntual en cuanto a la modificación en la secuencia del DNA, como punto de partida para su influencia sobre la estructura y función proteicas.

Clases y tipos de mutaciones puntuales y medianas, y su grado de incidencia		
Clase	Tipo	Incidencia
Sustitución		Tipo de mutación comparativamente común en DNA codificante y no codificante
	Transiciones y transversiones	Más comunes las transiciones que las transversiones, sobre todo en DNA mitocondrial
	Sustituciones sinónimas y no sinónimas	Las sinónimas son mucho más comunes que las no sinónimas en DNA codificante; las conservadoras son más comunes que las no conservadoras
	Sustitución de múltiples bases	Infrecuente, excepto en ciertos loci repetidos en tándem o repeticiones agrupadas
Inserción	De uno o unos pocos nucleótidos	Muy común en DNA no codificante pero infrecuente en DNA codificante
	Expansiones repetidas de tripletes	Infrecuente, pero puede contribuir a varios trastornos, especialmente neurológicos
	Otras inserciones grandes	Infrecuente; produce ocasionalmente duplicaciones en tándem e inserción de elementos transponibles
Eliminación o deleción	De uno o unos pocos nucleótidos	Muy común en DNA no codificante pero infrecuente en DNA codificante
	Grandes deleciones	Infrecuente; ocurre a menudo en regiones que contienen repeticiones en tándem o entre repeticiones dispersas

Las mutaciones puntuales son tan representativas que a veces se asumen como el único tipo de mutación existente. Es importante resaltar que es la **estabilidad** del cambio lo que hace que la mutación se perpetúe en la progenie celular, al pasar por replicación a cada nueva copia y permanecer en las secuencias de DNA de cada individuo. Estas variaciones genéticas heredadas contribuyen a las diferencias entre individuos dentro de una población. En algunos casos, cuando la modificación se produce en una región génica (estructural o reguladora) puede ser suficiente para alterar su mensaje, afectando a la funcionalidad de su producto génico y conduciendo a una enfermedad hereditaria. Al depender de la alteración de un único gen, estas enfermedades se llaman *monogénicas*, y se caracterizan por transmitirse a la descendencia según las leyes de Mendel. Como excepción, las mutaciones en el genoma mitocondrial se transmiten por herencia materna exclusivamente, debido a que el número de mitocondrias que aporta el espermatozoide en la formación del cigoto es muy inferior al procedente del óvulo.

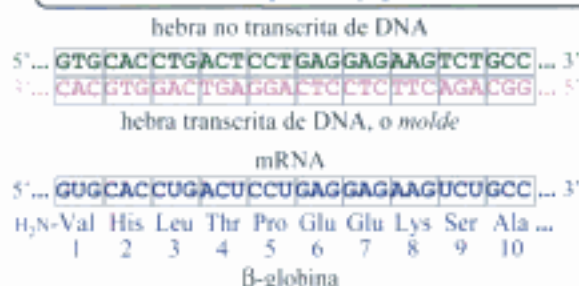
25.4.1 Cambios en la secuencia del DNA

La descripción de los distintos tipos de mutaciones se hará atendiendo a su causa a nivel molecular en el DNA: mutaciones por sustitución, por inserción y por delección.



Este estudio se acompaña, en cada uno de sus apartados, de ejemplos específicos indicativos de las consecuencias patológicas de la mutación.

Para la descripción de las mutaciones puntuales y sus efectos sobre la proteína codificada se hará uso, siempre que sea posible, de la secuencia de 30 pb de DNA que codifica los primeros 10 aminoácidos de la proteína β -globina humana (con un total de 147 aminoácidos):



La secuencia del mRNA resultante de la transcripción es igual que la de la hebra de DNA no-molde, no transcrita (salvo el cambio de T por U)

25.4.1.1 Sustitución

Como su nombre indica, la mutación consiste en la aparición de un nucleótido (coloquialmente, una base) en una posición de la secuencia ocupada originalmente por otro; dicho de otro modo, un **cambio** de una sola base (o nucleótido). Es relativamente común en el DNA, tanto codificante como no codificante. En casos raros, pueden reemplazarse simultáneamente varias bases agrupadas.

Según la naturaleza purínica o pirimidínica de las dos bases implicadas, se distingue entre transición y transversión.

a) Transiciones

Son sustituciones de una pirimidina por otra (C $\rightarrow$ T; T $\rightarrow$ C) o de una purina por otra (A $\rightarrow$ G o G $\rightarrow$ A). Existen, pues, 4 posibilidades, una para cada base original:

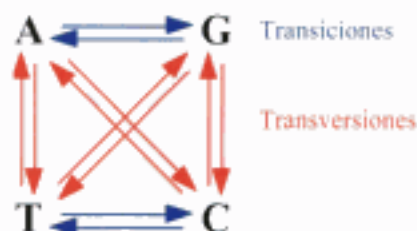
transición	par inicial	par final
A $\rightarrow$ G	A/T	G/C
G $\rightarrow$ A	G/C	A/T
T $\rightarrow$ C	T/A	C/G
C $\rightarrow$ T	C/G	T/A

b) Transversiones

Son sustituciones de una purina por una pirimidina (A o G $\rightarrow$ T o C), o viceversa (T o C $\rightarrow$ A o G). Existen 8 posibilidades, dos por cada base original:

transversión	par inicial	par final
A $\rightarrow$ T	A/T	T/A
A $\rightarrow$ C	A/T	C/G
G $\rightarrow$ T	G/C	T/A
G $\rightarrow$ C	G/C	C/G
T $\rightarrow$ A	T/A	A/T
T $\rightarrow$ G	T/A	G/C
C $\rightarrow$ A	C/G	A/T
C $\rightarrow$ G	C/G	G/C

Teóricamente, si la sustitución de las bases fuera aleatoria, la frecuencia de una transversión en el genoma debería ser mayor que la de una transición, pues hay doble posibilidad de que ocurra. Sin embargo, se observa en diferentes procesos mutagénicos que las transiciones son más comunes que las transversiones, especialmente en el DNA no funcional.



Como ejemplo específico puede citarse la transición de una base (G→A) en la región 5' no traducida del gen del factor IX de la coagulación (región que contiene islotes CpG, págs. 167 y 201). El cambio de secuencia en esa región reguladora impide la unión de un factor de transcripción clave que, en condiciones normales, activa el promotor del gen. Como consecuencia, disminuye la expresión del gen y se sintetiza menos factor IX, lo que produce hemofilia de tipo B, trastorno de la coagulación con una actividad coagulante disminuida en dos tercios.

25.1.1.1 Deleción, pérdida o eliminación

Consiste en la pérdida de uno o más nucleótidos de una secuencia. Es muy común en el DNA no codificante, pero infrecuente en el DNA codificante. Normalmente son mutaciones pequeñas, y sólo en raras ocasiones afectan a zonas lo suficientemente grandes como para ser visibles a nivel citogenético (en regiones con repeticiones en tándem o entre repeticiones dispersas). Pueden originar un cambio en el marco de lectura (pág. 350).

25.1.1.2 Inserción

Se trata de la situación contraria a la anterior, la aparición de uno o varios nucleótidos adicionales en una secuencia. Es muy común en el DNA no codificante, pero rara en el DNA codificante. Un ejemplo es la "repeticón de trinucleótidos" o "expansión de tripletes", un aumento en el número de copias de un triplete de nucleótidos observado en un gen o en sus proximidades.

25.1.2 Efecto sobre la secuencia de la proteína sintetizada

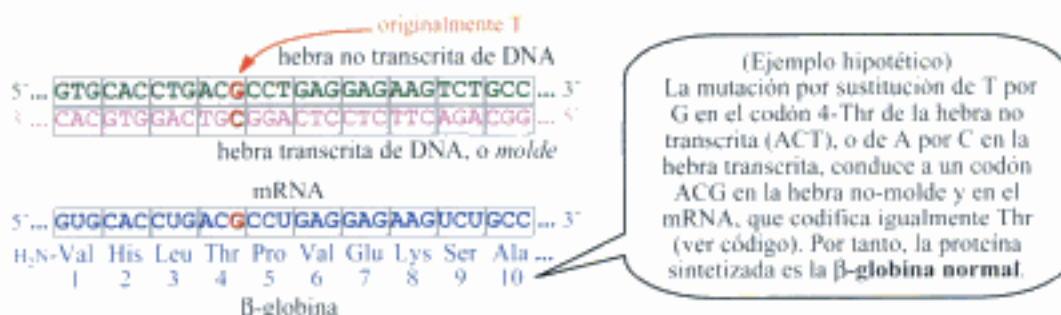
El cambio de secuencia del DNA que supone la mutación puede tener distintos efectos sobre la secuencia aminoacídica codificada y, en consecuencia, sobre la secuencia de la proteína que se sintetice y su función. Se distingue por ello entre mutaciones *silenciosas* y *no silenciosas*.

Tipo de mutación	Causa posible		
	Sustitución	Deleción	Inserción
Mutaciones silenciosas	✓		
Mutaciones no silenciosas	✓	✓	✓
De sentido alterado	✓		
Conservadoras			
No conservadoras			
Desplazamiento del marco		✓	✓
Sin desplazamiento del marco		✓ (triple)	✓ (triple)
Terminación prematura	✓	✓	✓
Terminación retrasada	✓	✓	✓

(cuando no se indica lo contrario, el cambio afecta a una sola base)

25.1.2.1 Mutaciones silenciosas

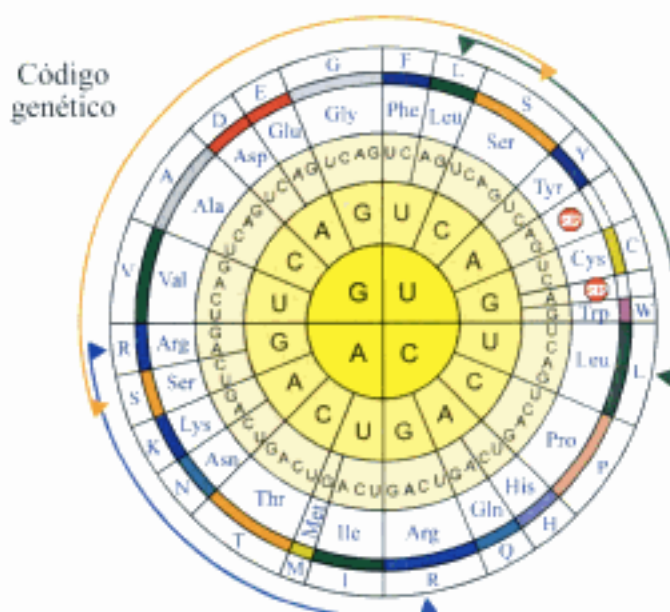
Se las llama también **sinónimas**, **neutrales**, **asintomáticas** o **con sentido** (del inglés *sense mutations*), y son muy numerosas (23-25% de todas las mutaciones en DNA codificante). Se caracterizan porque, a pesar de haber un cambio en la secuencia del DNA, no se altera la secuencia de su producto proteico. Esto es posible debido a la degeneración del código genético (varios tripletes o codones codifican un mismo aminoácido, pág. 295). Normalmente, la mutación es silenciosa porque afecta a la 3ª base de un codón de tal forma que el nuevo triplete es sinónimo del original, es decir, codifica el mismo aminoácido.



Igual ocurre si la sustitución de la base ocurre en la 1ª posición de determinados tripletes; concretamente, CUA $\leftrightarrow$ UUA y CUG $\leftrightarrow$ UUG, que codifican Leu (ver código), y AGA $\leftrightarrow$ CGA y AGG $\leftrightarrow$ CGG, para Arg.

Obviamente, las mutaciones silenciosas pasan inadvertidas, pues no se detectan a nivel fenotípico ni confieren ventaja o desventaja alguna al organismo en cuyo genoma surgen. En consecuencia, no están sometidas a la presión de selección evolutiva ni influyen en la susceptibilidad frente a enfermedades. Sin embargo, pueden originar una enorme diversidad genética en una población, conocida como **polimorfismo genético**, en este caso no génico (tema 26).

(se incluye aquí una copia del código genético para facilitar el análisis de los ejemplos)



25.4.2.2 Mutaciones no silenciosas

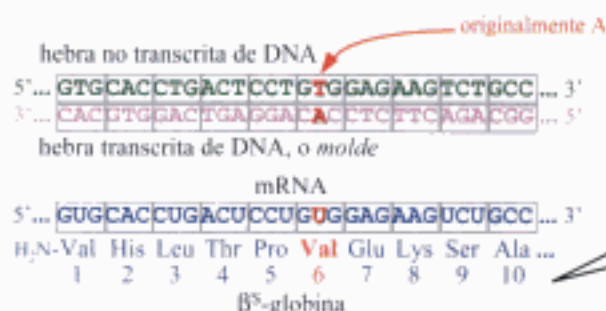
En este caso, la alteración en la secuencia de nucleótidos afecta a la secuencia proteica, pues codifica uno o varios aminoácidos diferentes de la secuencia original. Dependiendo del efecto que tenga la alteración de la secuencia proteica sobre su función, se puede distinguir entre mutaciones con efecto beneficioso y con efecto perjudicial. Aunque unas pocas mutaciones (menos del 1%) pueden tener un **efecto beneficioso o positivo**, por ejemplo favoreciendo la función de la proteína codificada o la interacción entre genes, son tan poco frecuentes que, en la práctica, es común que no se consideren. En cualquier caso, los mecanismos que las generan son los mismos que si el efecto es perjudicial. Por el contrario, interesa estudiar con todo detalle las mutaciones no silenciosas que se manifiestan con **efecto negativo, deletéreo o catastrófico**.

A continuación, se comentan los efectos más frecuentes de las mutaciones no silenciosas sobre los codones y, en consecuencia, sobre la secuencia del producto proteico.

a) Mutaciones que alteran el sentido

Contrariamente a las mutaciones silenciosas, donde la sustitución de una base origina un nuevo triplete que codifica el mismo aminoácido, es más común que dicho cambio produzca un codón nuevo que codifica un aminoácido diferente. Se las llama por ello mutaciones **no sinónimas, de sentido equivocado, falso o alterado** (en inglés, *missense mutation*). La mayoría de las mutaciones en regiones codificantes son de este tipo (73%), por lo que son causa de muchos trastornos. Además, la formación de más de una versión proteica supone que el **polimorfismo genético** generado por estas mutaciones da lugar también a un **polimorfismo proteico** (tema 26).

Lógicamente, estas mutaciones son particularmente relevantes cuando tienen lugar en **regiones génicas estructurales**. Por ejemplo, son de este tipo la mayoría de las mutaciones que producen hemoglobinopatías.



La sustitución de A por T en el codón 6-Glu de la hebra no transcrita (GAG), o de T por A en la hebra transcrita, conduce a un codón GUG en el mRNA, que codifica Val en lugar de Glu. Esta sola mutación (hemoglobina S, α₂β_S) es la causa de la **anemia falciforme** (págs. 382 y 396).

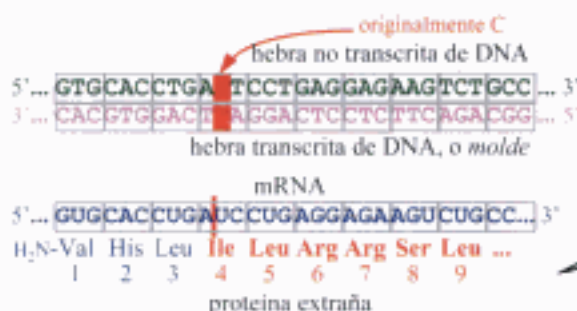
Dentro de las mutaciones de sentido alterado se distinguen dos subtipos:

- **Sustitución conservadora.** El aminoácido es sustituido por otro con una cadena lateral químicamente similar, por lo que el efecto sobre la estructura y función de la proteína puede ser mínimo. Esta situación se aproxima, pues, a las mutaciones silenciosas. En cierto modo, parece que el código genético se ha adaptado para reducir el efecto de las mutaciones, pues, además de la degeneración, los codones que especifican aminoácidos similares son a su vez parecidos. Por ejemplo, los codones para Asp (GAC, GAU) y Glu (GAA, GAG) son muy parecidos, de modo que el cambio de la 3ª base en un codón GAN (donde N es cualquier nucleótido) o bien es silencioso o tiene un efecto mínimo, el cambio de un aminoácido ácido por el otro. También pueden ser conservadores algunos cambios en la primera posición del codón, por ej. CUN (Leu) ↔ GUN (Val), donde N representa cualquier nucleótido.
- **Sustitución no conservadora.** El aminoácido es sustituido por otro no similar, diferente en carga, polaridad, tamaño de la cadena lateral, etc. Son no conservadoras casi todas las sustituciones en la 1ª o 2ª posición de un codón; por ejemplo, CGN (Arg) → GGN (Gly), CCN (Pro), CUN (Leu) o CAN (Gln/His), etc.

b) Mutaciones que cambian el marco de lectura

Este es un concepto importante, por la frecuencia con la que afecta a la secuencia del producto proteico. El **desplazamiento, desfase o cambio del marco de lectura** (*frameshift mutation*) consiste en que la inserción o delección de nucleótidos, aun sin afectar en gran medida a la secuencia de bases, cambia la forma como se leen los tripletes (pág. 288), por lo que produce una alteración importante en la secuencia de la proteína.

Se produce un cambio de marco cuando hay inserciones o delecciones de uno o más pares de bases, en número no múltiplo de 3 (es decir, un número no entero de codones). El cambio en la secuencia de aminoácidos se produce a partir de la primera inserción o delección, hasta el extremo carboxilo-terminal de la proteína. Por ello, en general la proteína resulta muy alterada en su estructura.



(Ejemplo hipotético)
La pérdida o delección de C en el codón 4-Thr de la hebra no transcrita (ACT), o de G en la hebra transcrita, cambia este triplete y todos los siguientes (se dice que se ha desplazado el marco de lectura), apareciendo una secuencia totalmente alterada en la globina.

Como ejemplo específico puede citarse la mutación más común encontrada en judíos ashkenazi (es decir, con antepasados de Europa del Este) que padecen la **enfermedad de Tay-Sachs**, un trastorno autosómico recesivo con acúmulo de gangliosidos en los lisosomas neuronales. El gen afectado es el que codifica la subunidad α de la enzima lisosómica **hexosaminidasa A** (isoenzima A de la β-N-acetilhexosaminidasa). La deficiencia enzimática total es responsable del fenotipo grave observado clásicamente en la forma infantil de la enfermedad.



La inserción de 4 nucleótidos (TATC) en la hebra no transcrita, o de ATAG en la hebra transcrita, cambia el marco de lectura. Además de cambiar todos los aminoácidos de la hexosaminidasa a partir del punto de inserción, en el nuevo marco hay un codón de terminación más adelante, por lo que la proteína es más corta y no funcional.

c) *Mutaciones que no cambian el marco*

La inserción o deleción de 3 pb en el DNA (o múltiplos de tres), aunque añade o elimina algún aminoácido a la proteína, no cambia el marco de lectura, por lo que el resto de la secuencia aminoacídica es normal. El efecto de estas mutaciones es, por tanto, mucho menor y puede no afectar a la función de la proteína.

Como ejemplo específico, en el 70% de todos los alelos asociados con la **mucoviscidosis**, o **fibrosis quística**, se observa una deleción triple, en el cromosoma **7q**. La mutación afecta al gen de una proteína de membrana, transportadora de aniones cloruro. La proteína mutada se pliega incorrectamente, por lo que durante su procesamiento postraducciona queda retenida en el lumen del retículo endoplásmico, unida a proteínas carabina (págs. 337-338); se impide así su llegada a la membrana plasmática, con consecuencias que afectan a varios tejidos y glándulas endocrinas (insuficiencia pancreática y afecciones respiratorias). La enfermedad se diagnostica por un exceso de cloruro en el sudor.



La pérdida de 3 nucleótidos (CTT) en la hebra no transcrita, o de GAA en la hebra transcrita, produce un acortamiento en la cadena peptídica, pero al ser la deleción un múltiplo de 3 nucleótidos, se recupera el marco de lectura. El codón 507 se ha visto afectado por la mutación, pero, casualmente, debido a la degeneración del código, no se altera el aminoácido (AUC:Ile → AUU:Ile). El residuo Phe-508 desaparece, por lo que la proteína pierde un aminoácido; la mutación se denomina $\Delta F508$ (deleción de la fenilalanina 508).

d) *Mutaciones con terminación prematura de la proteína*

Algunas mutaciones no afectan a la secuencia de aminoácidos codificados, sino que la sustitución, la inserción o la deleción de uno o varios nucleótidos tiene como efecto la aparición de un **codón de terminación** o **de paro** (UAA, UAG o UGA en el mRNA citoplásmico de eucariotas superiores, pág. 294). Como consecuencia, se provoca el cese prematuro de la traducción del mRNA y se obtiene una proteína acortada (en su región C-terminal).

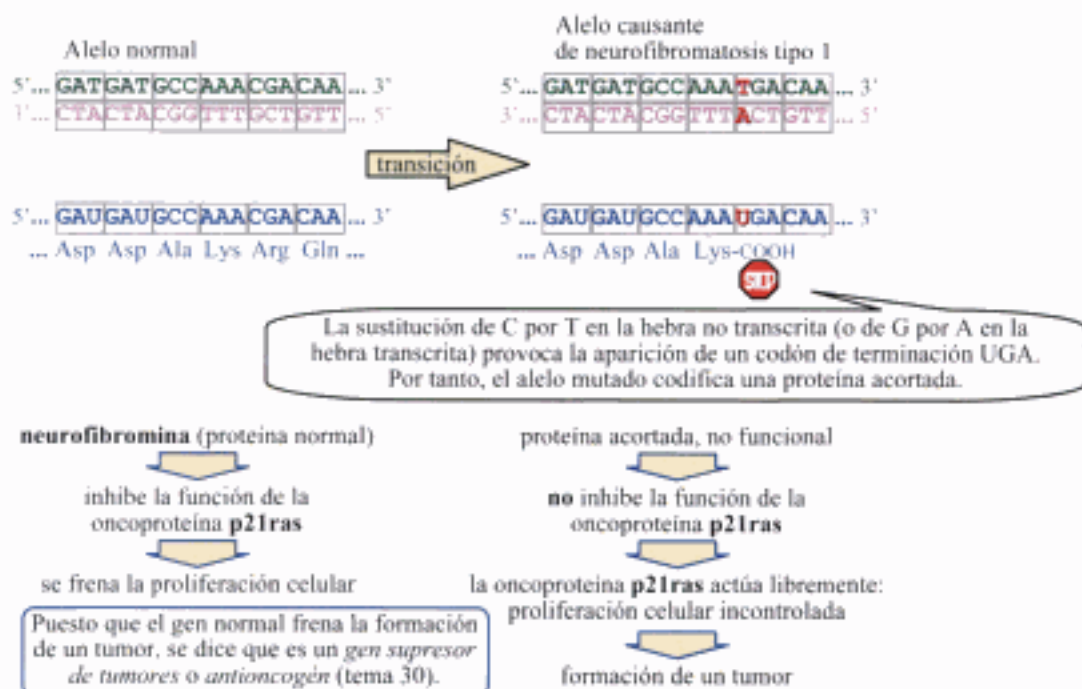
Este tipo de mutaciones, también llamadas **mutaciones sin sentido** (*nonsense mutations*, que no debe confundirse con *missense mutations*, mutaciones de sentido equivocado por sustitución de un aminoácido por otro) son poco frecuentes; por ejemplo, suponen cerca del 4% de las sustituciones en el DNA codificante. El acortamiento de la proteína es a menudo considerable, por lo que se ven afectadas dramáticamente su estructura tridimensional, su estabilidad y su función. Con frecuencia esta alteración acelera la degradación intracelular de la proteína no funcional, ocasionando generalmente un efecto fenotípico.



(Ejemplo hipotético)

La sustitución de A por T en el codón 8-Lys de la hebra no transcrita (TAG) (o de T por A en la hebra transcrita) conduce a un codón UGA en el mRNA, que es un codón de paro, de terminación o codón sin sentido. La síntesis de la proteína se detiene al llegar a él, por lo que la proteína resulta acortada o truncada.

Como ejemplo específico, la mutación causante de la **neurofibromatosis tipo 1 (NF1)** o **enfermedad de Recklinghausen**, un trastorno autosómico dominante (pág. 395) localizado en el cromosoma **17q**.



La mutación puede producir directamente la aparición del codón de terminación, como en los ejemplos mostrados, pero también puede hacerlo indirectamente, mediante un cambio en el marco de lectura (véase, como ejemplo, el polimorfismo de grupo sanguíneo ABO, pág. 369).

e) *Mutaciones con terminación retrasada*

De la misma forma que una mutación puntual puede provocar la aparición de un codón de terminación, también puede eliminar uno existente, sustituyéndolo por otro que codifica un aminoácido. En ese caso, la traducción del mRNA continúa más allá del extremo 3' codificante original, hasta que se alcance un nuevo codón de terminación. Por tanto, se sintetiza una proteína de mayor longitud, cuya secuencia de aminoácidos en la región C-terminal no tiene significado estructural ni funcional. Las consecuencias fenotípicas son, en general, similares a las de la terminación prematura.

25.5 CAUSAS Y MECANISMOS BÁSICOS DE LAS MUTACIONES

El DNA está sometido a daños, lesiones y modificaciones de diversos tipos, provocados por distintas causas. Éstas pueden ser fortuitas o endógenas y provocadas, inducidas o exógenas.

25.5.1 Mutaciones endógenas

Son mutaciones que se generan por situaciones o agentes propios del ambiente intracelular; bajo condiciones normales, éstos son la fuente mayoritaria de mutaciones. Pueden originarse por errores en la replicación, o bien por reacciones que ocurren de forma espontánea como consecuencia de la inestabilidad química de la molécula de DNA o de la acción de subproductos del metabolismo celular (mutágenos endógenos).

25.5.1.1 Incorporación de nucleótidos erróneos durante la replicación

La replicación por la DNA polimerasa, moviéndose a lo largo del DNA, combina la incorporación de nucleótidos a la hebra nueva con la corrección de pruebas, de forma que duplica la doble hélice con una gran fidelidad. Más del 99,9% de los errores de replicación son rápidamente corregidos por enzimas reparadoras que reconocen el nucleótido incorrecto en la hebra que se está sintetizando y lo reemplazan por uno correcto, complementario al de la hebra molde. Los errores de replicación observados ocurren con una frecuencia muy baja, de 10^{-9} a 10^{-11} mutaciones por nucleótido incorporado en cada ciclo de división celular. Esta baja tasa global de mutación significa que sólo se introduce un nucleótido incorrecto por cada 10.000 Mb. Puesto que el genoma diploide humano tiene 6.600 Mb, ello se traduce en **menos de una nueva mutación por cada división celular**. A pesar de ello, el gran número de divisiones celulares durante la vida de un individuo hace que el número de mutaciones sea significativo.

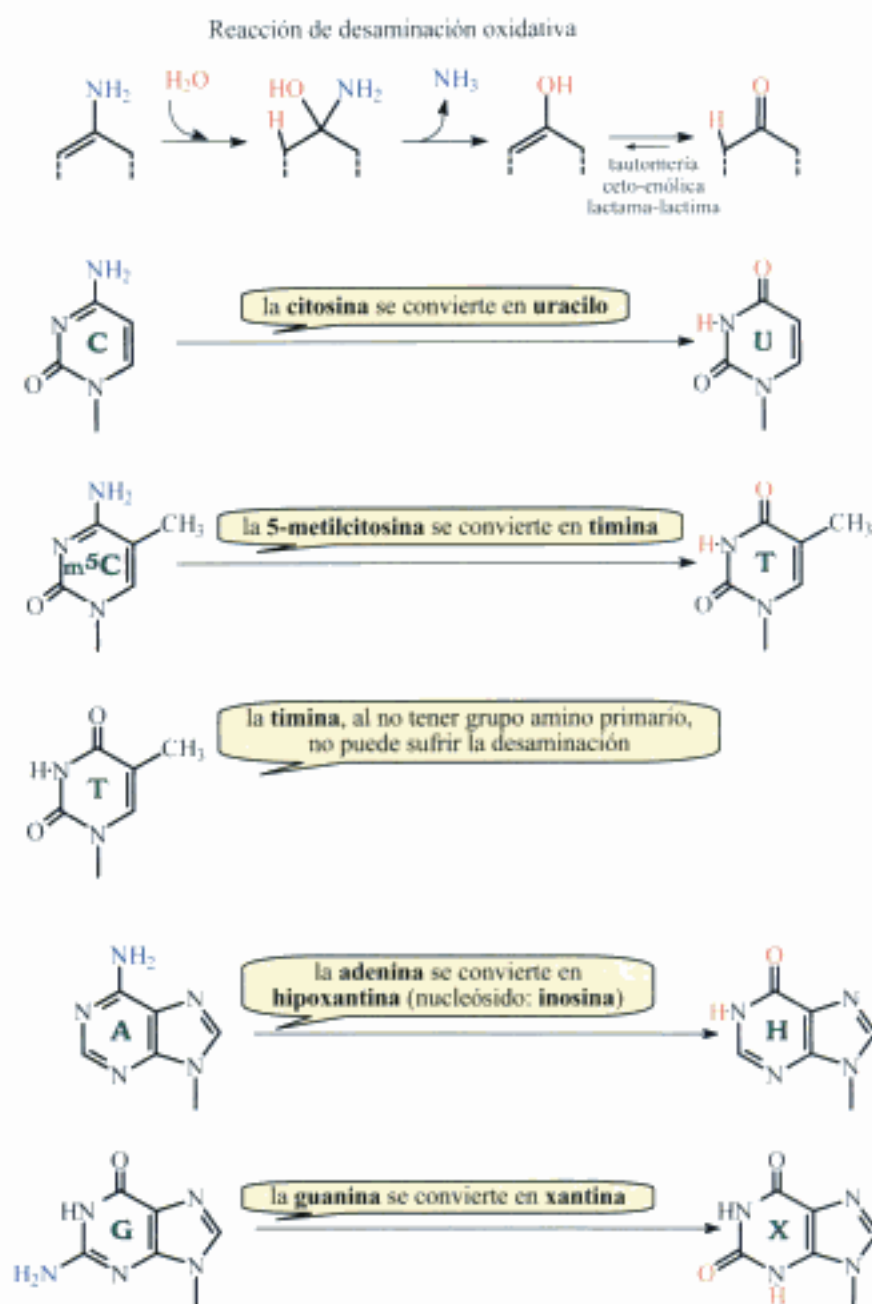
Una de las causas de incorporación errónea de nucleótidos está en la tautomería de las bases nitrogenadas (pág. 14); las formas tautoméricas menos estables (imina y lactima) establecen puentes de hidrógeno diferentes, lo que permite el apareamiento de un nucleótido incorrecto en la hebra de nueva síntesis. Además de los errores en la identidad del nucleótido incorporado, durante la replicación también se introducen mutaciones debido a deslizamientos de la DNApol sobre la hebra molde, que producen deleciones o inserciones por repetición de una pequeña secuencia.

25.5.1.2 Otras mutaciones espontáneas

Las bases en el DNA pueden experimentar espontáneamente reacciones (sin intervención de enzima alguna) que transforman su identidad, de modo que si no son corregidas por los sistemas de reparación se convierten en una mutación.

a) Sustituciones por desaminación oxidativa

Consisten en la pérdida de un grupo amino exocíclico, con aparición de un grupo carbonilo anular. Ello da lugar a que se altere la formación de puentes de hidrógeno entre las bases.

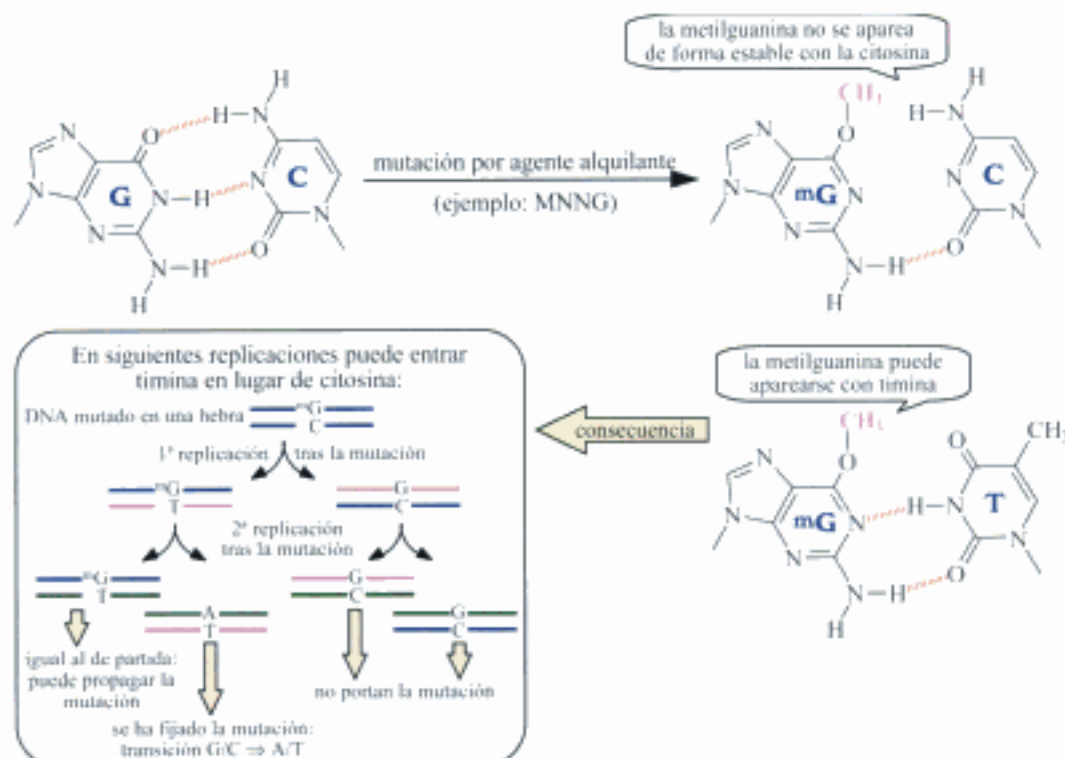
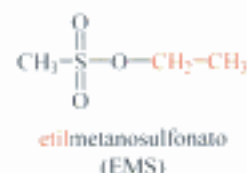
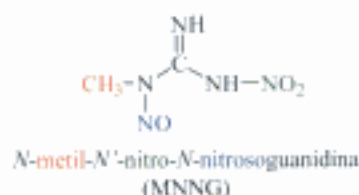


Las transformaciones C→U y 5-metil-C→T se producen lentamente, pero unas 100 veces más rápidamente que las A→H y G→X. Por ejemplo, una desaminación de citosina al día por cada 10^7 C en el DNA de una célula típica de mamífero (lo que equivale a unas 100 mutaciones por día).

• Agentes químicos alquilantes

Transfieren grupos, generalmente metilo o etilo, a un átomo nucleofílico (O o N) de las bases nitrogenadas. Entre los mejor estudiados están:

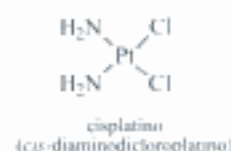
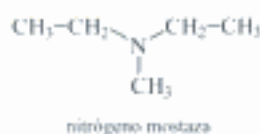
- Dimetilsulfato: $(\text{CH}_3)_2\text{SO}_4$
- Dimetilnitrosamina: $(\text{CH}_3)_2\text{N}_2\text{O}$
- Metilnitrosourea
- Nitrosoguanidinas
- Etilmetanosulfonato



• Agentes intercalantes

Al insertarse entre los pares de bases apilados del DNA, distorsionan su estructura helicoidal. Algunos provocan el desplazamiento del marco de lectura, mientras que otros (agentes bifuncionales) establecen enlaces cruzados entre bases de la misma hebra o de hebras opuestas.

- Naranja de acridina, proflavina, etc.
- Benzo[a]pireno, uno de los productos que se encuentran en el humo.
- Nitrógeno mostaza
- Compuestos de cisplatino



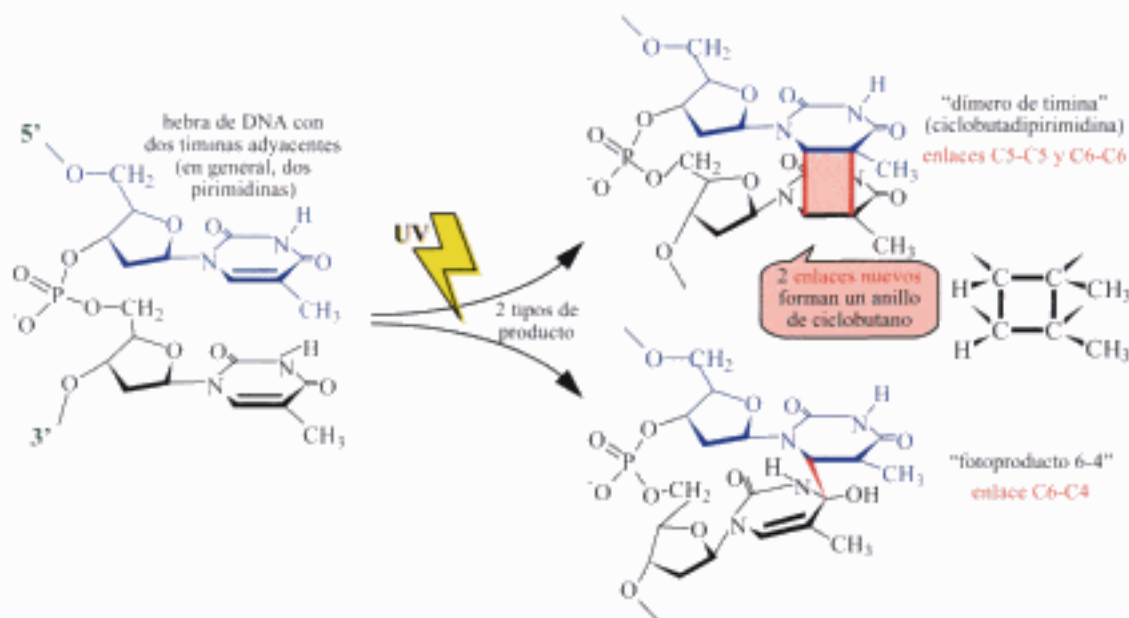
• Otros agentes químicos

- 5-bromouracilo: se incorpora en lugar de T durante la replicación, gracias a su analogía estructural (pág. 132).
- Ácido nitroso (HNO_2): acelera la desaminación oxidativa C→U.
- Compuestos que pueden metabolizarse a ácido nitroso, como nitritos o nitratos.
- Hidrocarburos aromáticos policíclicos: benzo[a]pireno, 2-acetamidofluoreno, N-metil-4-antraceno.
- Aminas aromáticas: N-metil-4-aminoazobenceno (MAB).
- Varias drogas: ciclofosfamida, dietilestilbestrol.
- Compuestos naturales: dactinomicina, aflatoxina.
- Compuestos inorgánicos: arsénico, asbestos, berilio, cadmio, cromo.

25.5.2.2 Lesiones y mutaciones inducidas por agentes físicos

Son varios los agentes que originan lesiones por exposición, con efectos mutagénicos o cancerígenos.

La **radiación ultravioleta**, componente de la luz solar, provoca la reacción entre dos residuos de timidina adyacentes en la misma hebra, formando a partir de sus dobles enlaces anulares un sistema tricíclico con un anillo de ciclobutano, que se conoce como **dímero de timina**. Esta asociación covalente de dos T impide que puedan aparearse con sus A complementarias y origina una estructura tridimensional localmente incompatible con el giro de la doble hélice, que forma un "codo" en la molécula de DNA. Además, se paraliza la replicación, pues la DNA polimerasa no puede reconocer los dímeros y se detiene. La irradiación UV también origina otra reacción entre T adyacentes, con formación de dímeros diferentes llamados fotoproductos, en los que el enlace aparece entre el C-6 de una T y el C-4 de la adyacente.



Las **radiaciones ionizantes** (rayos X y rayos γ) pueden provocar apertura de los anillos, fragmentación de las bases, así como rotura del esqueleto covalente del DNA. También generan (por fotólisis del agua) radicales hidroxilo, que producen 8-hidroxiguanina, entre otros. Debido a su mayor poder de penetración, estas radiaciones afectan a todo tipo de tejidos, a diferencia de la luz UV, que provoca lesiones solamente en la piel.

25.6 REPARACIÓN DEL DNA

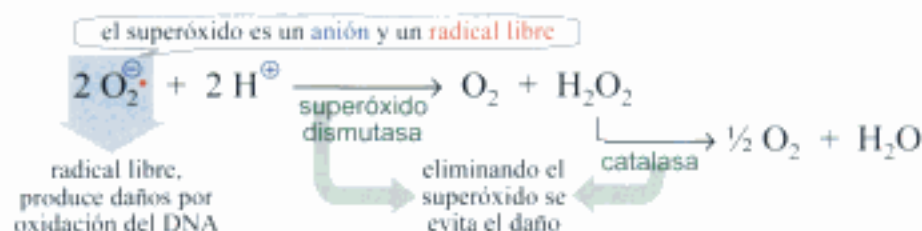
La capacidad de una célula para detectar y reparar daños en el DNA es esencial para el mantenimiento de la integridad estructural de su genoma en general y de sus genes en particular. Existen con este objeto diversos sistemas enzimáticos *in vivo*, que pueden llegar a ser tan complejos como el propio aparato de replicación, lo que pone de manifiesto su gran importancia para la supervivencia celular. De hecho, la velocidad de mutación es el reflejo del balance entre el número de errores ocurridos y corregidos: una baja tasa de mutación indica una gran eficiencia de los sistemas de reparación, mientras que un fallo de éstos conduce a un aumento del número de mutaciones.

En general, no puede establecerse una estrategia única de reparación por parte de la célula, sino que existen distintas vías.

Tipo de reparación	Ejemplo	Tipo de lesión reparada
Eliminación de mutágenos (destoxificación)	Superóxido dismutasa	Se evita la formación de la lesión oxidativa
Reversión directa	Fotoliasa	Fotodímeros de pirimidina
	Alquiltransferasas	Derivados alquilo de las bases
Escisión de nucleótidos (general)	Sistema de la escisinucleasa	Las que causan distorsión en la doble hélice (fotoproductos y sustituyentes voluminosos)
Escisión de bases (específica)	N-glicosilasas del DNA y endonucleasas AP	Bases desaminadas, modificadas, oxidadas, etc. y sitios abásicos
Reparación de apareamientos incorrectos		Nucleótidos mal apareados por errores en la replicación o bases modificadas

25.6.1 Eliminación de agentes mutágenos

El primer mecanismo para reducir el daño al DNA no es propiamente una vía de reparación, sino que consiste en evitar las lesiones antes de que se produzcan, mediante la neutralización de compuestos mutágenos. Uno de tales mecanismos de detoxificación lo ejerce la superóxido dismutasa, que evita la acumulación de radicales superóxido, agentes causantes de lesiones oxidativas en el DNA.

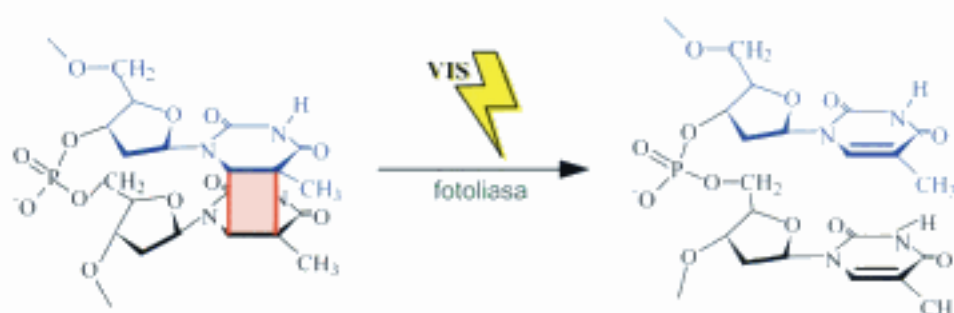


25.6.2 Reversión directa de la lesión

Es la forma más inmediata de reparar una mutación, la regeneración de la base dañada revirtiendo la transformación química que ha sufrido. Este mecanismo no es factible en la mayoría de los casos, a excepción de los ejemplos siguientes.

25.6.2.1 Reparación de fotodímeros

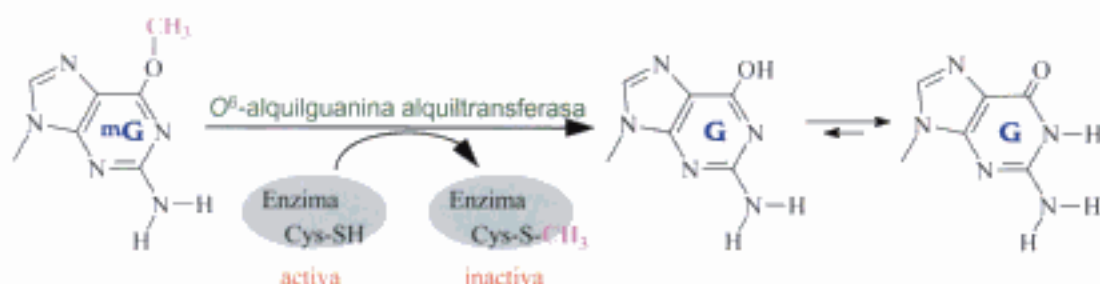
Ya se ha explicado que la luz ultravioleta provoca la formación de anillos de ciclobutano entre dos pirimidinas adyacentes (pág. 356), lo que supone una distorsión estructural que impide físicamente la replicación y la transcripción. El mecanismo de reparación se basa en el reconocimiento del fotodímero por la enzima **fotoliasa**, que lo escinde, mediante una reacción de transferencia electrónica iniciada por absorción de luz visible, regenerando sus dos bases originales. La fotoliasa tiene como coenzimas FADH_2 y un cromóforo que absorbe la luz e inicia la transferencia de electrones. Este requerimiento de luz (azul o violeta) hace que se la llame también **enzima fotorreactivadora**. En condiciones de oscuridad se requieren otros sistemas para reparar estas lesiones. Este sistema se ha encontrado en algunas especies, pero no en otras; así, existe en bacterias, hongos, la levadura *Saccharomyces cerevisiae*, peces, reptiles, anfibios y mamíferos marsupiales, pero no aparece en *Schizosaccharomyces pombe* ni en mamíferos placentarios, incluyendo a los humanos. En algunos organismos, incluyendo *Drosophila*, se ha detectado otra fotoliasa que repara los fotoproductos 6-4, ausente a su vez en *E. coli*, levaduras y humanos.



25.6.2.2 Reversión de bases modificadas con grupos alquilo

El caso más conocido, muy bien estudiado en *E. coli* e igualmente presente en eucariotas (levaduras y humanos) y en todas las especies estudiadas, es la reparación de O^6 -metilguanina y O^6 -etilguanina, productos de la modificación del DNA por agentes mutágenos alquilantes, como nitrosoguanidinas, metilnitrosourea y etilmetanosulfonato (pág. 355).

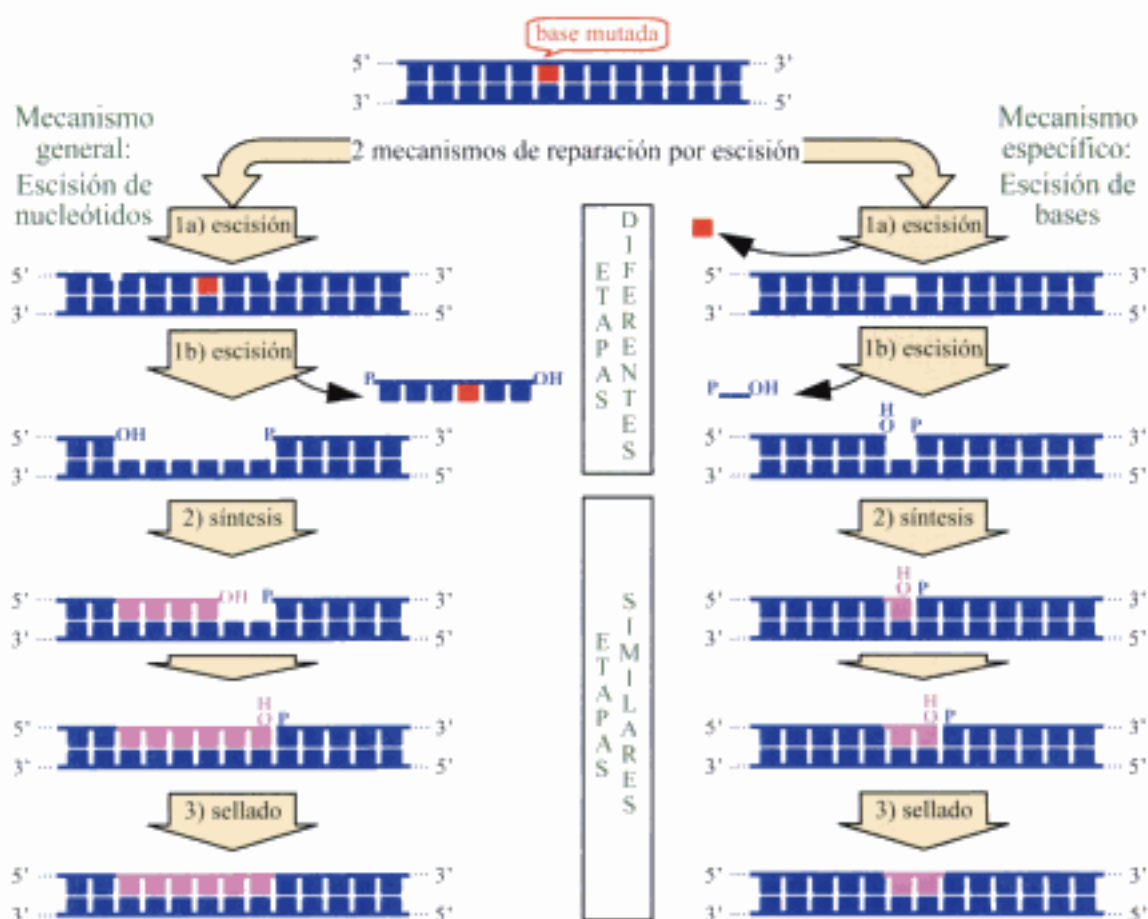
El grupo alquilo de la base es eliminado por una alquiltransferasa, que lo transfiere a un residuo cisteína de su propio centro activo. Éste no puede regenerarse, por lo que la "catálisis" inactiva la propia enzima (con lo que estrictamente no es enzima ni catalizador, pues no se cumple la condición de que no se altere como consecuencia de la reacción que acelera). Dada esta inactivación, la célula debe sintetizar nuevas moléculas de alquiltransferasa para mantener el proceso reparador, por lo cual en presencia de un nivel excesivo de mutaciones alquilantes puede tener lugar una "saturación" del sistema enzimático de reparación.



25.6.3 Reparación por escisión

Se conoce bajo este nombre genérico a la reparación de lesiones basada en la eliminación de la base nitrogenada, nucleótido o segmento de DNA defectuosos y su reemplazo por otro nuevo, de secuencia correcta. A diferencia de la reparación por reversión, éste es un proceso complejo, que tiene lugar en varias fases, que utiliza varias proteínas (sólo algunas específicamente reparadoras) y que conlleva siempre la síntesis de un fragmento de DNA (denominada *síntesis fuera de programa*). El proceso se inicia a través del reconocimiento molecular de la lesión (base alterada o cambio en la estructura tridimensional del DNA).

De forma global, en el mecanismo de reparación se pueden distinguir 3 etapas, la primera distinta para los mecanismos "general" y "específico", y las dos restantes comunes:

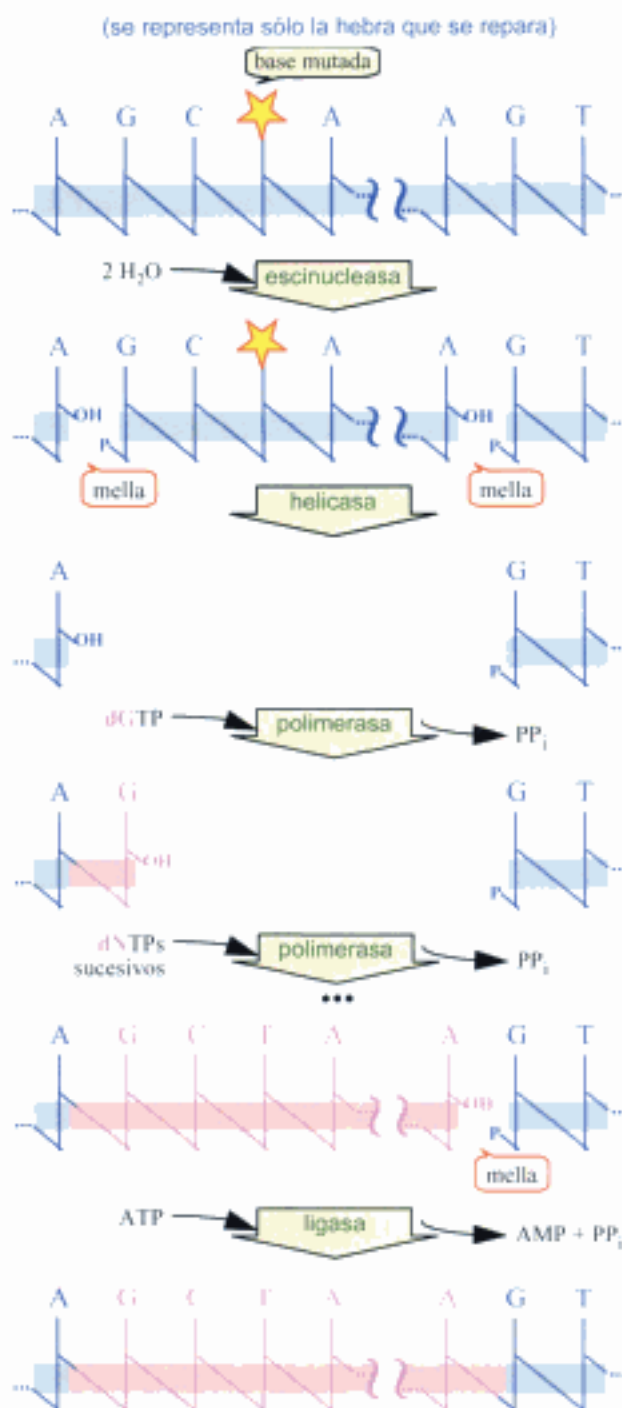


25.6.3.1 Mecanismo general de reparación por escisión de nucleótidos

En este mecanismo (**NER**, *nucleotide excision repair*) intervienen proteínas que reconocen distintos tipos de alteraciones que causan distorsión en la doble hélice, como los dímeros de timina o una base con un sustituyente voluminoso (psoraleno, cisplatino, acetamidofluoreno...). En humanos, es el sistema de defensa más importante contra el daño al DNA causado por el humo del tabaco, agente responsable de más del 30% de las muertes por cáncer en el mundo.

- El proceso comienza con el **reconocimiento de la lesión** por unión de proteínas componentes del complejo de reparación, con gasto de ATP.

- En la **fase de escisión** un complejo proteico que incluye endonucleasas hidroliza sendos enlaces fosfodiéster de la hebra lesionada a ambos lados del punto de mutación (pero no junto a él; concretamente, a -22 y +6 en humanos). La peculiaridad de cortar sólo una de las hebras y en dos puntos separados la distingue de otras nucleasas, por lo que se la denomina **escinucleasa**; en humanos está formada por 16 polipéptidos. A continuación, una **helicasa** favorece la separación del oligonucleótido liberado (de 12 ó 13 nt en procariontes y unos 28 nt en eucariotas). Algunas de las subunidades de la escinucleasa permanecen unidas al tramo de hebra sencilla hasta que sean desplazadas por la maquinaria de replicación.



- A continuación, en la **fase de "síntesis de reparación"**, una **DNA polimerasa** rellena el hueco dejado por la escinucleasa, mediante la elongación del extremo 3'-OH formado en el corte anterior. El uso de la otra hebra como molde asegura que se introduce la secuencia correcta. En eucariotas, realizan esta síntesis las polimerasas δ y ϵ , junto con sus factores accesorios **PCNA** y **RFC** (pág. 154).

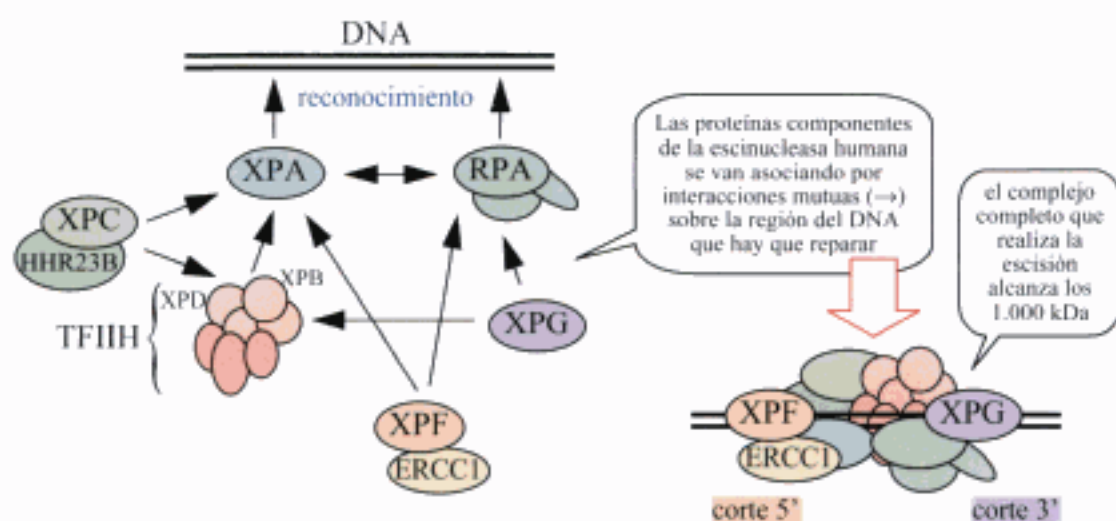
- Finalmente, una vez que la polimerasa ha rellenado todo el hueco y sólo queda una mella, una **DNA ligasa** une el extremo 3'-OH recién sintetizado con el extremo 5'-P formado en el corte de la hebra original, cerrando la mella (**fase de sellado o ligamiento**).

Este mecanismo de reparación posee la particularidad de estar relacionado con la transcripción, gracias a que participan proteínas comunes (principalmente, el factor de transcripción TFIIH, pág. 267). La RNAPol II detenida en el punto de lesión puede actuar de señal para el complejo de reparación, y además se observa una reparación preferente de las regiones de DNA que se transcriben y, dentro de ellas, de la hebra molde.

La escinucleasa humana es un gran complejo proteico (cerca de 1.000 kDa en total) que se puede considerar formado por 6 subunidades funcionales que se van asociando y separando a lo largo del proceso, constituidas por un total de 16 cadenas polipeptídicas. Se conocen ya prácticamente todos sus genes, las proteínas y su función. Si se suman las proteínas responsables de la replicación y sellado, en total participan en el proceso de reparación por escisión de nucleótidos unos 30 polipéptidos diferentes.

Subunidad	Nº de cadenas polipeptídicas	Función																					
XPA	1	Reconoce la lesión en el DNA y se une a él. Pose un dedo de zinc C ₄ (pág. 63).																					
RPA	3	Reconoce la lesión en el DNA y se une a él. Es la misma proteína de unión a hebra sencilla que interviene en la replicación.																					
TFIIH	7	Es el mismo factor de transcripción general de los genes de clase II. Actúa por su actividad helicasa, separando las hebras del DNA en la región a reparar. Su actividad proteína quinasa no es necesaria. <table border="1" data-bbox="566 659 1347 873"> <tr> <td>Helicasa:</td><td>XPB</td><td>Helicasa 3'→5'</td></tr> <tr> <td></td><td>XPB</td><td>Helicasa 5'→3'</td></tr> <tr> <td></td><td>p62</td><td></td></tr> <tr> <td></td><td>p44</td><td></td></tr> <tr> <td>Proteína quinasa:</td><td>p41</td><td>Quinasa dependiente de ciclinas CDK7</td></tr> <tr> <td></td><td>p38</td><td>Ciclina H</td></tr> <tr> <td></td><td>p34</td><td></td></tr> </table>	Helicasa:	XPB	Helicasa 3'→5'		XPB	Helicasa 5'→3'		p62			p44		Proteína quinasa:	p41	Quinasa dependiente de ciclinas CDK7		p38	Ciclina H		p34	
Helicasa:	XPB	Helicasa 3'→5'																					
	XPB	Helicasa 5'→3'																					
	p62																						
	p44																						
Proteína quinasa:	p41	Quinasa dependiente de ciclinas CDK7																					
	p38	Ciclina H																					
	p34																						
XPC+HHR23B	2	Función desconocida, posiblemente ayuda a estabilizar el complejo escinucleasa																					
XPG	1	Endonucleasa que corta el extremo 3' de la región escindida. Presenta homología de secuencia con la endonucleasa FEN1 (pág. 156)																					
XPF+ERCC1	2	Endonucleasa que corta el extremo 5' de la región escindida.																					

Nota: los nombres XPA-XPG proceden de xerodermia pigmentosa (pág. 363)



25.6.3.2 Mecanismos específicos de reparación por escisión de bases

Este sistema (**BER**, *base excision repair*) se encarga de reparar bases específicas; por tanto, reconoce y actúa en sitios concretos, reparando lesiones que son demasiado sutiles para distorsionar gravemente el DNA, es decir, no son reconocidas por el sistema general de escisión de nucleótidos. Por ejemplo, puede corregir daños por hidrólisis (desaminación o pérdida de base), oxidación por radicales libres de oxígeno, alquilación, y derivados no voluminosos como la timina glicol, N³-metiladenina, 8-hidroxiguanina, etc.

(se representa sólo la hebra que se repara)

- La **fase de escisión** se diferencia de la reparación general porque, en primer lugar, se elimina la base alterada o incorrecta, y no su nucleótido completo. Una **N-glicosilasa de DNA** reconoce un solo tipo de base, anómala o apareada incorrectamente, e hidroliza su enlace N-glicosídico con la desoxirribosa, dando lugar a un **sitio apirimidínico o apurínico**, según sea el caso.

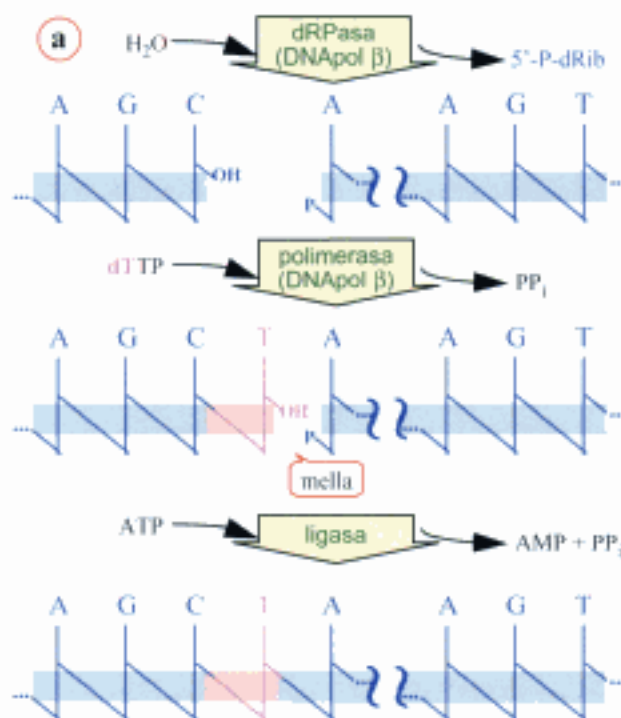
En todas las células existen varias N-glicosilasas, específicas para bases distintas y que, por tanto, reparan lesiones diferentes.

- Una vez generado el sitio abásico, en una **segunda subetapa de la fase de escisión** se produce el corte de la hebra de DNA junto a esa desoxirribosa que no está unida a una base nitrogenada.

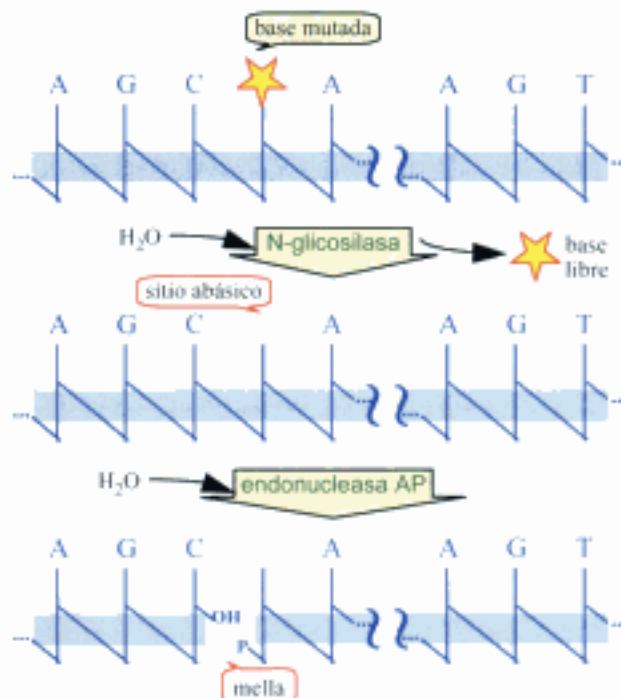
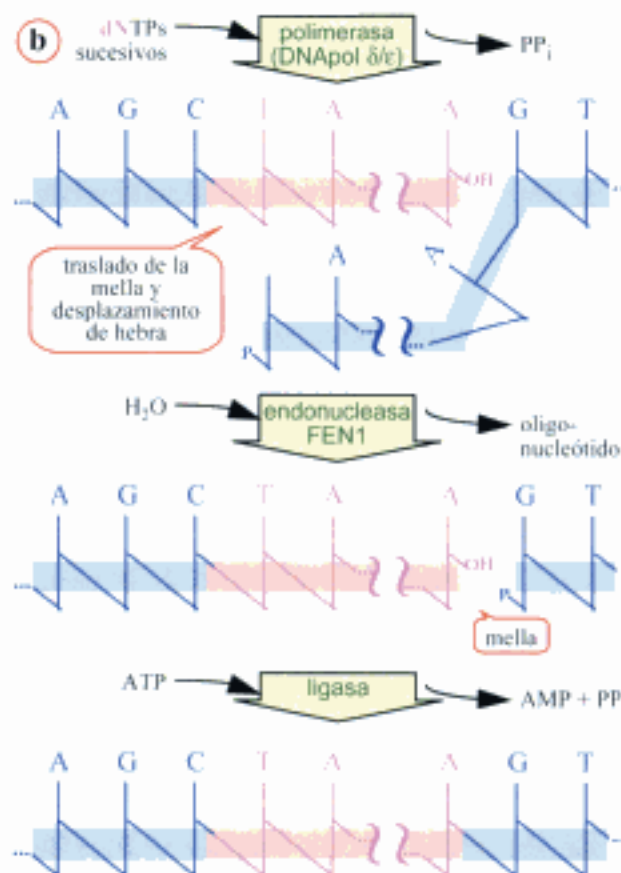
Entran en juego para ello las **endonucleasas AP** (de **apurínico** y **apirimidínico**), que hidrolizan un enlace fosfodiéster, originando una mella; en eucariotas (levaduras y mamíferos) cortan siempre al lado 5' del sitio abásico, generando un extremo nucleótido-3'-OH y otro 5'-fosfodesoxirribosa.

- A partir de este momento, la reparación continúa de forma similar al mecanismo general, gracias a la intervención de una **DNA polimerasa** que elonga sustituyendo el nucleótido abásico y, en ocasiones, varios más (e introduce la secuencia correcta por complementariedad con la otra hebra). En eucariotas se han encontrado dos mecanismos:

a) Reparación de un solo nucleótido, mediante la DNApol β , que posee dos actividades enzimáticas: **desoxirribosofosfodiesterasa**, que elimina el nucleótido abásico, y **polimerasa**, que lo sustituye por uno completo. Finalmente, una **ligasa** completa la reparación cerrando la mella.



b) Reparación de un tramo más largo, realizada por las **polimerasas δ o ϵ** , junto con sus factores accesorios PCNA y RFC, y la **endonucleasa FEN1** (págs. 154 y 156). Finaliza igualmente con el sellado de la mella por una ligasa.



Como ejemplo de N-glicosilasa, la **uracilo-DNA glicosilasa** es una enzima muy específica, que no actúa sobre residuos U del RNA, pero que recorre el DNA reconociendo como extraño el uracilo (que sólo se diferencia de la timina en la carencia de grupo metilo). En las células se reduce la posibilidad de que U se incorpore en la síntesis del DNA gracias a una **dUTP pirofosfatasa** que elimina el nucleótido precursor ($UTP \rightarrow UMP + PP_i$). Sin embargo, en el DNA ya sintetizado se forma U con gran frecuencia por la desaminación espontánea de C (pág. 353). Si el DNA estuviese formado por U (en lugar de T) sería imposible diferenciar los residuos U formados a partir de C, y la mutación U→C no podría ser reparada. Gracias a que el componente del DNA es T, y no U, es posible la reparación de la mutación frecuente U→C y el mantenimiento del par A/T. Este mecanismo se considera una de las ventajas por las que el DNA (con A, G, T, C) ha podido evolucionar como un mejor depositario de la información genética. Sin embargo, no ocurre lo mismo con las citosinas metiladas, relativamente comunes en el DNA eucariótico, cuya desaminación oxidativa produce T, que no puede ser distinguido de los T propios del DNA.

Otras glicosilasas de DNA reparan de forma análoga diversas bases modificadas:

Enzima	Actúa sobre	Lesión que repara
uracilo-DNA glicosilasa	U	desaminación de C
3-metiladenina-DNA glicosilasas	3-Me-A, 3-Me-G, 7-Me-G, O ² -Me-C, O ² -Me-T, hipoxantina	metilación de A, G, C o T, desaminación de A
formamidopirimidina-DNA glicosilasa	8-hidroxi-guanina, formamidopirimidinas	oxidación de G
hidrato de pirimidina-DNA glicosilasa	timina glicol, pirimidinas oxidadas o fragmentadas	oxidación de pirimidinas

(*) Me = metil

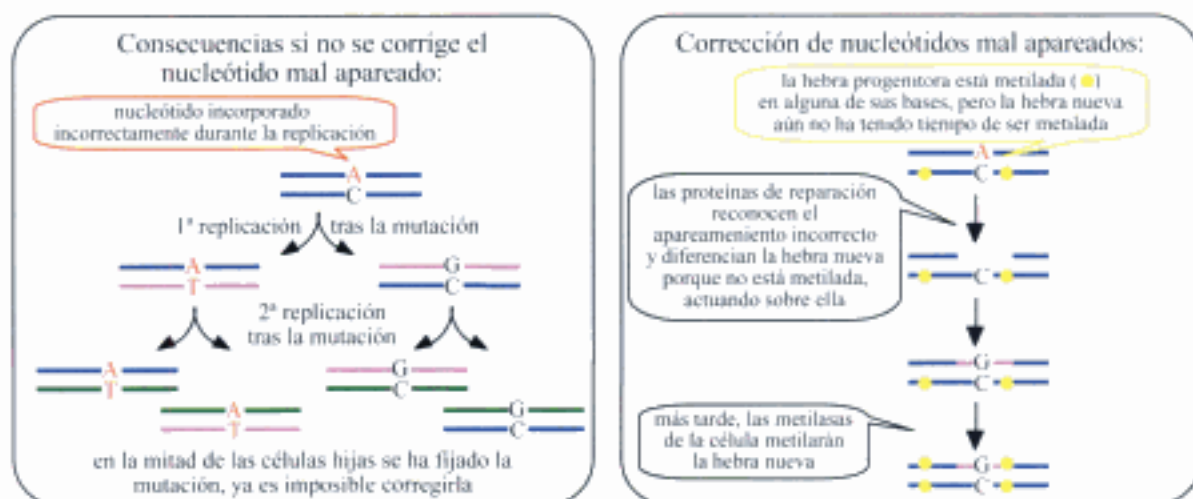
Finalmente, debe mencionarse que el reconocimiento de los sitios AP es vital no sólo para completar la acción previa de las N-glicosilasas, sino para corregir la despurinación espontánea, que es un hecho relativamente frecuente (pág. 354).

25.6.4 Reparación de apareamientos incorrectos

Normalmente, los mecanismos de reparación anteriores actúan al poco tiempo de que se ha producido la mutación, antes de que ésta se perpetúe debido a la replicación. Sin embargo, existen además otros procesos, denominados de **reparación posreplicación**, que corrigen errores cometidos en la replicación o errores que la impiden (actúan, por tanto, una vez por cada división celular). Entre ellos se encuentran la **reparación por recombinación** y la **reparación de apareamientos incorrectos**. Comentaremos solamente esta última.

Los apareamientos incorrectos se producen por errores de la replicación o como consecuencia de recombinación; el mismo sistema es también capaz de corregir pequeñas inserciones de 1 a 4 nucleótidos en una de las hebras. Normalmente, la mayoría de los errores cometidos inicialmente por la DNA polimerasa se rectifican gracias a su propia actividad correctora de pruebas, o exonucleasa 3', pero unos pocos pueden permanecer sin corregir. Es entonces cuando entra en juego este mecanismo, formado por proteínas que **reconocen** la presencia de un par de bases incorrectamente apareadas y **eliminan** la base errónea o bien un tramo de la hebra de DNA que la contiene. Éstas son luego **repuestas** por los mecanismos de reparación ya estudiados, con intervención de las DNA polimerasas δ y ϵ , y sus cofactores RFC y PCNA. En humanos se han podido identificar 5 genes que codifican proteínas de este sistema de reparación, llamados *hMSH2*, *hMLH1*, *hPMS1*, *hPMS2* y *GTBP/hMSH6*.

La cuestión importante que surge es: ¿cómo determina la maquinaria de reparación cuál de las dos bases de un par incorrecto es la errónea, la que hay que eliminar? Parecen existir varios mecanismos para esto, pero uno de ellos se basa en la **metilación** del DNA (pág. 272). Las enzimas encargadas de la metilación de bases tardan un tiempo en actuar, por lo que en el DNA dúplex recién replicado la hebra nueva aún no está metilada, mientras que la hebra progenitora sí lo está. Esto permite a las proteínas reparadoras diferenciar la hebra nueva, que ha de ser la que contiene la mutación, y corregir ésta. Este mecanismo se conoce bien en *E. coli*, y posiblemente actúe bajo principios similares en eucariotas.



25.7 ENFERMEDADES HUMANAS ASOCIADAS A LA REPARACIÓN

Se conoce una serie de enfermedades atribuidas a defectos en genes que codifican proteínas de reparación del DNA. Estos defectos conducen con frecuencia a una mayor incidencia del cáncer, pues se produce una acumulación de mutaciones no corregidas.

El cáncer de piel denominado **xerodermia pigmentosa, o pigmentaria (XP)**; también, xerodermia de Kaposi o melanosis lenticular progresiva), es una enfermedad hereditaria causada por una deficiencia en alguno de los genes responsables de la reparación por escisión de nucleótidos, conduciendo principalmente a la acumulación de dímeros de timina. La piel de estos enfermos es extremadamente sensible a la exposición a los rayos UV de la luz solar, ocasionándoles cánceres de piel con una frecuencia 1.000 ó 2.000 veces superior a la de personas normales. También sufren una incidencia de 10 a 20 veces mayor de cánceres en órganos internos, y síntomas neurológicos como retraso mental, ataxia progresiva y sordera. Se han identificado defectos causantes de XP en 7 genes del sistema de reparación, que codifican componentes de la escisinucleasa: *XPA*, *XPB*, *XPC*, *XPD*, *XPF*, *ERCC1* y *XPG* (pág. 360).

El **síndrome de Cockayne (CS)**, caracterizado por enanismo, retraso mental y síntomas neurológicos progresivos causados por desmielinización, también con una sensibilidad moderada a la radiación ultravioleta, se origina por mutaciones en los genes *CSA* o *CSB*, relacionados también con la reparación pero más bien con su conexión con la transcripción. Algunas mutaciones de *XPB*, *XPD* y *XPG* dan lugar a síndromes solapantes XP/CS.

Otra enfermedad genética relacionada con la reparación del DNA es la **tricotiodistrofia (TTD)**, debida a mutaciones en los genes *TTD-A*, *XPB* o *XPD*. Se trata más bien de un defecto de la transcripción, que se puede corregir *in vitro* por microinyección de TFIIH a las células. Sus síntomas (en los dos últimos casos, solapantes con XP) incluyen caries dental, ictiosis, anomalías esqueléticas y retraso mental progresivo por desmielinización.

El **cáncer colorrectal hereditario sin poliposis (HNPCC, hereditary nonpolyposis colorectal cancer)**, uno de los cánceres hereditarios más comunes, está causado por defectos en los genes *hMSH2* o *hMLH1*, ambos responsables de la reparación de apareamientos incorrectos. Se pueden detectar mediante VNTR, pues el sistema de reparación defectuoso no elimina pequeños bucles de hebra desapareada de DNA, lo que conduce a una mayor longitud de algunos microsatélites de DNA (pág. 383).

Otras enfermedades asociadas a defectos en los sistemas de reparación son la **ataxia-telangiectasia**, asociada con una mayor sensibilidad al daño del DNA por rayos X, y la **anemia de Falconi**, que parece deberse a un defecto en la reparación del daño por entrecruzamiento químico de bases.

Tema 26

Diversidad del genoma: polimorfismos

26.1	Introducción	365
26.2	Diversidad genética dentro de una especie	366
26.3	Mecanismos implicados en la variabilidad genética	366
26.3.1	Duplicación del material genético	367
26.3.2	Transposición	367
26.4	Consecuencias funcionales del polimorfismo	368
26.4.1	Polimorfismo en regiones codificantes, o polimorfismo génico	368
26.4.1.1	Sin efecto fenotípico	368
26.4.1.2	Con alteración del fenotipo	369
26.4.2	Ejemplos representativos de polimorfismo génico	369
26.4.2.1	Polimorfismo de grupo sanguíneo ABO	369
a)	Identidad de los grupos sanguíneos y aspectos clínicos	369
b)	Genética y bioquímica del sistema ABO	370
26.4.2.2	Polimorfismo en el locus de la galactosemia	373
a)	Descripción y bioquímica de la galactosemia	373
b)	Genética de la galactosemia	374
26.4.2.3	Polimorfismo de la α_1 -antitripsina	376
26.4.3	Polimorfismo en regiones génicas no codificantes	376
26.4.4	Polimorfismo en regiones no génicas	376
26.4.5	Polimorfismo en el genoma mitocondrial	377

26.1 INTRODUCCIÓN

El DNA genómico nuclear es una entidad extremadamente variable, no sólo entre individuos de especies diferentes (variación máxima, interespecie), sino también entre individuos de una misma especie (variación mínima, intraespecie). Esta diversidad es responsable de fenómenos biológicos de tanta trascendencia como la evolución de las especies y, dentro de cada especie, de la presencia de características diferenciales en cada individuo, únicas e irrepetibles. La **diversidad o variabilidad genética** se debe a variaciones en la secuencia del genoma; por tanto, en un sentido amplio el concepto de **diversidad** se hace sinónimo de **polimorfismo** (en su significado literal, *muchas formas*) o, más concretamente, polimorfismo genético, cromosómico, de secuencia o de DNA. Los proyectos de secuenciación del genoma ya tienen en cuenta esta diversidad; normalmente, se secuencia el genoma de individuos concretos, representativos de distintas razas o etnias (pág. 240).

El polimorfismo afecta tanto a regiones codificantes del genoma (**polimorfismo génico**) como no codificantes (**polimorfismo genético** en general). En ambos casos puede consistir en la variación de un solo par de bases del DNA o, menos frecuentemente, de millones de pb (macromutación). El primer caso se conoce hoy día como polimorfismo de un solo nucleótido o **SNP** (acrónimo inglés de *single-nucleotide polymorphism*; coloquialmente, *"snip"*). Cualquier variación puede tener lugar en células germinales o reproductoras, con lo que se transmite a la descendencia (polimorfismo hereditario, por ejemplo en las enfermedades congénitas), o bien en células somáticas, no reproductoras, en cuyo caso no se transmite (polimorfismo no hereditario, por ejemplo en la mayoría de los cánceres).

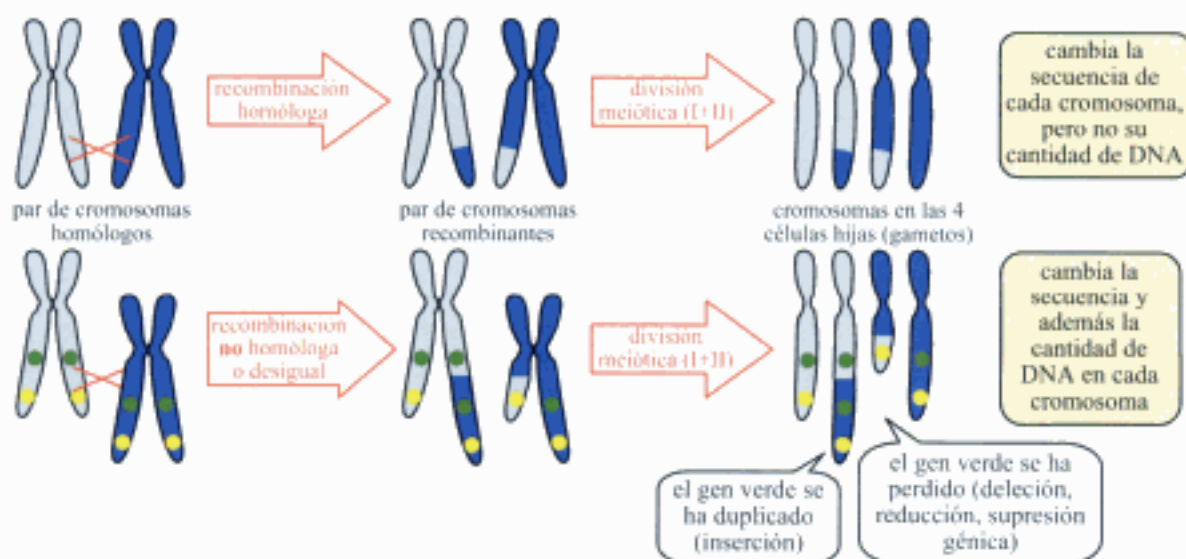
Inicialmente, la definición de polimorfismo se refería a las proteínas, y luego a sus genes. Hoy día, se debe definir como la "existencia simultánea en una población de genomas con distintos alelos para un locus determinado", sea éste génico o no. Los **alelos** son variaciones de la secuencia de DNA presente en un **locus** o posición definida en un cromosoma (pág. 94); consecuentemente, en una célula diploide cada locus autosómico está ocupado por dos alelos, uno de origen paterno y otro materno, situados en sendos cromosomas homólogos. El **grado de polimorfismo** de una población es

26.3.1 Duplicación del material genético

Se habla de duplicación **a gran escala** cuando afecta a grandes porciones de un cromosoma. Sin entrar a describir los aspectos concretos del proceso (propios de una Genética Molecular), debe indicarse que ha actuado progresivamente durante la evolución, mediante una variedad de mecanismos complejos.

La duplicación **a pequeña escala** es de interés cuando tiene lugar en regiones codificantes del genoma, afectando a genes específicos, por lo que también se llama duplicación génica, selectiva de genes específicos. La aparición de múltiples copias del gen, inicialmente idénticas, suele evolucionar posteriormente hacia la formación de copias ligeramente diferentes. De esta forma, una copia ancestral de un gen puede originar todos los miembros de una **familia multigénica** (genes que codifican polipéptidos con secuencias relacionadas, pág. 110). Estas familias codifican diversas proteínas típicas de eucariotas; por ejemplo, globinas (proteínas de transporte y almacén de oxígeno en sangre y músculo), actinas (parte del aparato contráctil en células musculares y del citoesqueleto en células no musculares), opsinas (que detectan la luz a diferentes longitudes de onda en células de la retina), colágeno (en diferentes tipos de tejido conjuntivo), albúminas e inmunoglobulinas (en plasma sanguíneo), etc.

El principal mecanismo responsable de la duplicación de genes completos a lo largo de la evolución y, en definitiva, de la generación de tan gran variedad y cantidad de DNA repetitivo en las regiones no codificantes y codificantes del genoma eucariótico, es la **recombinación no homóloga o desigual**. Ésta ocurre durante la meiosis, entre homólogos no correctamente alineados de un par de cromosomas, en los que, por tanto, no coinciden las regiones equivalentes (como sí ocurre en la recombinación homóloga), pero sí se enfrentan regiones con similitud de secuencia. Como consecuencia del sobrecruzamiento de cromátidas (pág. 106), un homólogo puede recibir mayor cantidad de material genético que el otro (lo que supone en ambos, respectivamente, una inserción o duplicación y una delección).



También han surgido por duplicación génica los genes que codifican, en tejidos diferentes, las **isoformas** (formas alternativas de una proteína) o, en su caso, las **isoenzimas** o **isozimas** (formas alternativas de una enzima). Por ejemplo, la **fosfatasa alcalina**, codificada por al menos 4 genes diferentes, que se expresan de forma diferencial en cada tejido. Tres de ellos están agrupados cerca del telómero del cromosoma 2q: ALPI de intestino, ALPP de placenta (con un 87% de similitud entre sí) y ALPPL, similar a la de placenta. El gen de la cuarta isozima, ALPL, de hígado, hueso, riñón y otros tejidos, está localizado cerca del telómero de 1p y muestra una similitud de secuencia de 57% con ALPI y 55% con ALPP. Tan elevada homología indica que estas formas específicas de la misma proteína enzimática se originaron por una duplicación génica ancestral.

26.3.2 Transposición

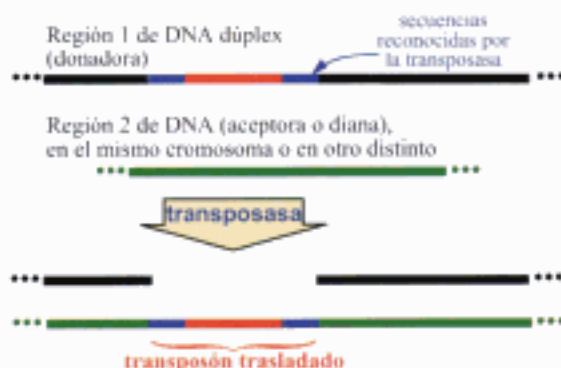
Otra fuente importante de variabilidad genética es el proceso de **transposición**, bien conocido en bacterias pero también presente en el genoma de las células eucariotas. Consiste en el desplazamiento, de un sitio a otro del cromosoma, de fragmentos de DNA llamados **elementos móviles**, **elementos transponibles** o **transposones**; su longitud es variada, desde unos cuantos cientos de nucleótidos a decenas de miles. La transposición es responsable en eucariotas de una gran proporción del DNA repetitivo no codificante, destacando las secuencias *Alu* y *LINE-1* (págs. 114-115), que por sí solas suponen el 10% del genoma humano.

Existen dos mecanismos de transposición en eucariotas:

En el **mecanismo de cortar y pegar**, que también se da en procariotas, el **transposón** se escinde del DNA donador para insertarse en el DNA diana, de modo que la molécula donadora queda interrumpida, cortada, y el transposón cambia de molécula. Los cortes y uniones están catalizados por enzimas específicas, **transposasas**, que

reconocen ciertas secuencias de los extremos del transposón. En algunos casos, la transposasa está codificada por parte de la propia secuencia del transposón.

DESPLAZAMIENTO DEL TRANSPOSÓN POR CORTE Y PEGADO



DESPLAZAMIENTO POR RETROTRANSPOSICIÓN



Mediante el **mecanismo de retrotransposición**, exclusivo de eucariotas, la región del DNA donador que se transpondrá se transcribe primero a RNA (por RNA polimerasas celulares). A partir de él, una transcriptasa inversa origina el correspondiente cDNA (págs. 195 y 206), que luego se inserta en el DNA diana. Este mecanismo es similar al de la replicación en retrovirus, de donde recibe el nombre. Debe destacarse que, a diferencia del mecanismo de cortar y pegar anterior, en este caso el transposón no cambia meramente de lugar, sino que además se duplica. Por último, algunos retrotransposones codifican en su propia secuencia la transcriptasa inversa necesaria para su propagación (por ejemplo, *LINE-1*).

26.4 CONSECUENCIAS FUNCIONALES DEL POLIMORFISMO

El polimorfismo puede tener distinta trascendencia desde el punto de vista funcional, dependiendo de si afecta a regiones no codificantes, reguladoras o codificantes del genoma y también del modo como afecte al mensaje genético de estas últimas.

26.4.1 Polimorfismo en regiones codificantes, o polimorfismo génico

De la parte codificante del genoma humano (alrededor de 100.000 genes), se cree que unas tres cuartas partes son **monomorfos**, genes únicos compartidos sin variación por todos los individuos. Éstos constituyen la parte del genoma común o "de referencia", que define la especie y determina sus características (por ej., en humanos, presencia de cabeza, tronco, extremidades, etc.). Tales genes son casi idénticos a los correspondiente en la especie más cercana (en nuestro caso, el gorila). El resto, una cuarta parte del total (unos 25.000 genes), son **polimorfos**, sus variaciones, en muchos casos mínimas, determinan el polimorfismo **génico**.

Dentro del polimorfismo génico, se pueden distinguir dos situaciones de acuerdo con su trascendencia funcional, es decir, su efecto sobre el fenotipo.

26.4.1.1 Sin efecto fenotípico

Aun estando la variabilidad localizada en una secuencia codificante, el polimorfismo puede no tener consecuencias sobre el fenotipo (es decir, sobre el producto génico), principalmente por dos motivos: que no se altere la secuencia proteica codificada (como consecuencia de la degeneración del código genético, caso de las *mutaciones silenciosas*, pág. 348) o que la variación ocurra en una región de la proteína que no es esencial para su función (entre otras, las *mutaciones conservadoras*, pág. 350). Podrían incluirse también en este último caso los polimorfismos que afectan a los intrones, regiones génicas pero no codificantes (que, por tanto, pueden considerarse también polimorfismos no génicos, pág. 376).

Este tipo de polimorfismo es el más común, responsable de una diversidad genética enteramente normal. Las proteínas formadas por los distintos alelos de un locus de este tipo son variantes relativamente comunes, con secuencias y estructuras idénticas o similares y manteniendo una funcionalidad similar. Aparece, por ejemplo, en proteínas plasmáticas como inmunoglobulinas, α_1 -antitripsina, haptoglobina, transferrina y ceruloplasmina, y en otras proteínas tanto intra- como extracelulares. Durante años, estas formas múltiples o variantes de proteínas fueron el principal producto apreciable del polimorfismo, gracias a su separación, principalmente mediante electroforesis.

26.4.1.2 Con alteración del fenotipo

En otros casos, la variación en la secuencia del gen modifica la secuencia de la proteína de tal forma que afecta a su funcionalidad. Aunque esta modificación puede ser beneficiosa (base de la evolución), lo más común es encontrar un polimorfismo patológico o deletéreo (perjudicial). La alteración puede modificar las características bioquímicas, fisiológicas o incluso morfológicas de la célula, pudiendo llegar a originar procesos patológicos, en función de la gravedad de dicho efecto fenotípico. Se puede distinguir entre:

- Variaciones fenotípicas que no influyen en la susceptibilidad a enfermedades, caso de los polimorfismos responsables de la estatura, el color de pelo, el grupo sanguíneo, etc.
- Variaciones fenotípicas que influyen de forma leve o mínima en la susceptibilidad a ciertos procesos patológicos, tales como los polimorfismos responsables de rasgos morfológicos complejos.
- Variaciones fenotípicas que desempeñan un papel importante en la aparición de una patología. Esto es lo que ocurre, por ejemplo, con las hemoglobinas anormales, que se manifiestan en forma de enfermedades genéticas. Precisamente, el concepto de **enfermedad genética** constituye la manifestación más obvia y a veces más extrema de un cambio genético.

26.4.2 Ejemplos representativos de polimorfismo génico

A continuación se van a desarrollar en detalle tres ejemplos de polimorfismo proteico bien conocidos, algunos con implicaciones patológicas, originados por sendos polimorfismos genéticos en regiones codificantes.

26.4.2.1 Polimorfismo de grupo sanguíneo ABO

a) Identidad de los grupos sanguíneos y aspectos clínicos

La existencia de los grupos sanguíneos humanos del sistema ABO fue el primer caso descrito (1990) de diversidad proteica determinada genéticamente. La existencia de tres alelos determina cuatro fenotipos en la sangre humana, denominados grupos sanguíneos O, A, B y AB. Éstos se deben a la presencia diferencial de dos antígenos, A y B, en la superficie de los eritrocitos. Estos antígenos son oligosacáridos diferentes entre sí, que forman parte de estructuras glucídicas mayores en glicoproteínas y glicolípidos, situados tanto en la cara externa de la membrana del eritrocito como de la mayoría de las células del organismo (especialmente epiteliales y endoteliales), así como en secreciones solubles (por ejemplo, en las mucinas, un tipo de glicoproteínas presente en saliva, jugo gástrico y secreción intestinal).

Una particularidad del sistema ABO es que en el plasma de una persona se encuentran habitualmente anticuerpos contra los antígenos de los que carecen sus células. Se trata de lo que se conoce como "anticuerpos que existen naturalmente" (*naturally occurring antibodies*, NAO), es decir, que no son el resultado de una reacción inmune. Son de tipo IgM (a diferencia de los anticuerpos inmunes, IgG) y empiezan a producirse a partir de los 3 ó 6 meses de vida. En consecuencia, en cada individuo existe una **relación recíproca** entre los antígenos presentes en los eritrocitos y los anticuerpos en el plasma.

Fenotipo	Grupo A	Grupo B	Grupo AB	Grupo O
Antígenos presentes en sus eritrocitos	 A	 B	 A y B	 ninguno
Anticuerpos naturales en el plasma	 anti-B	 anti-A	ninguno	 anti-A y anti-B

La trascendencia de estos anticuerpos naturales radica en que determinan la compatibilidad de transfusiones sanguíneas y de trasplante de tejidos. Una combinación compatible es aquella en la que no coinciden antígenos de los eritrocitos del donante con anticuerpos del plasma del receptor. En consecuencia, el grupo O es donante universal y el AB es receptor universal.

Compatibilidad en transfusiones

Los eritrocitos del donante no deben encontrar anticuerpos contra sus antígenos en la sangre del receptor (si así fuese, se produciría la agregación de los eritrocitos donados).

		Receptor				
		grupo A	grupo B	grupo AB	grupo O	
Donante		anti-B	anti-A	ninguno	anti-A y anti-B	anticuerpos que contienen
grupo A	A	+		+		
grupo B	B		+	+		
grupo AB	A y B			+		
grupo O	ninguno	+	+	+	+	

resultados (+ = compatible)

antígenos que poseen

Esta compatibilidad es también la base del método serológico de determinación del grupo sanguíneo:

Ensayo serológico para determinar el grupo sanguíneo ABO

Al añadir plasma o, más comúnmente, suero de un grupo sanguíneo conocido a los eritrocitos de una muestra de sangre, éstos forman agregados si el suero contiene anticuerpos contra los antígenos presentes en la superficie de los eritrocitos.

Muestra			Reactivo				anticuerpos que contienen
			Suero de				
			grupo A	grupo B	grupo AB	grupo O	
Eritrocitos de	grupo A	A		+		+	
	grupo B	B	+			+	
	grupo AB	A y B	+	+		+	
	grupo O	ninguno					

resultados
(+ = agregación)

antígenos
que poseen

b) Genética y bioquímica del sistema ABO

Desde el punto de vista de la biología molecular, es de sumo interés considerar las implicaciones en ambas áreas de un sistema polimórfico tan representativo como el de los grupos sanguíneos ABO.

El sistema ABO consiste, como se ha indicado, en un **polimorfismo de oligosacáridos** (carbohidratos), que es el resultado de un **polimorfismo proteico** en las glicosiltransferasas responsables de la síntesis de esos oligosacáridos. Éste se origina, a su vez, en un **polimorfismo génico**, la existencia de 3 alelos para el gen del sistema ABO que codifica la glicosiltransferasa. Por otra parte, la expresión de los antígenos A y B está condicionada además por la de otro gen polimórfico, el gen H, cuyo producto proteico es otra glicosiltransferasa que sintetiza el carbohidrato precursor (antígeno H) sobre el que actúan las transferasas codificadas por el gen ABO. Los aspectos particulares de todo el sistema se muestran comparativamente a continuación.

Genes que determinan los grupos sanguíneos A, B y O:

Gen H: codifica la glicosiltransferasa que genera el oligosacárido precursor o antígeno H.

Gen ABO: codifica la glicosiltransferasa que, a partir del antígeno H, genera los antígenos A o B.

Locus ABO:

en la banda cromosómica 9q34.1

El gen ABO cubre en total 19,5 kb, con 7 exones y 6 intrones.

Las regiones codificantes (exones) suman 1.062 pb.

Los alelos (A, B y O) se diferencian en varios nucleótidos de los exones 6 y 7.

Los productos génicos de los alelos A y B son polipéptidos de 354 aminoácidos, con una masa molecular de 41 kDa y actividades glicosiltransferasas distintas.

El producto del alelo O, con sólo 117 aminoácidos, no tiene actividad enzimática.

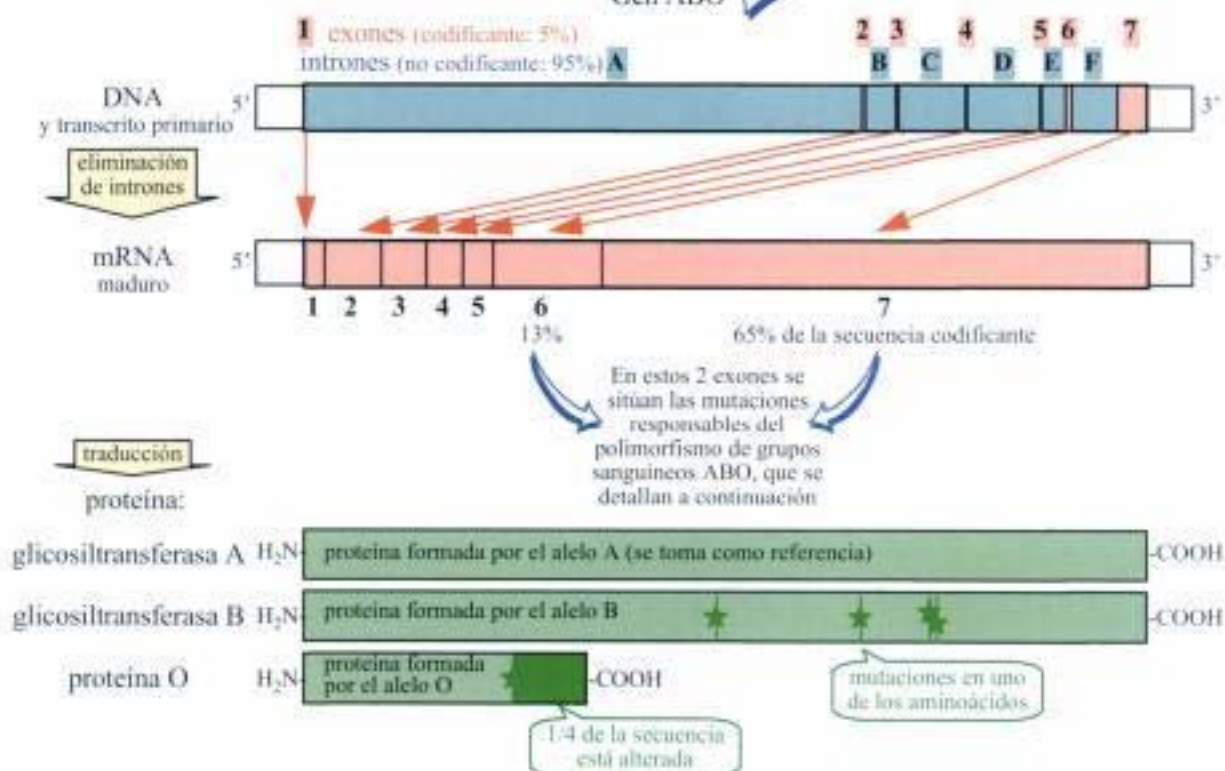
Locus H:

en el brazo cromosómico 19q

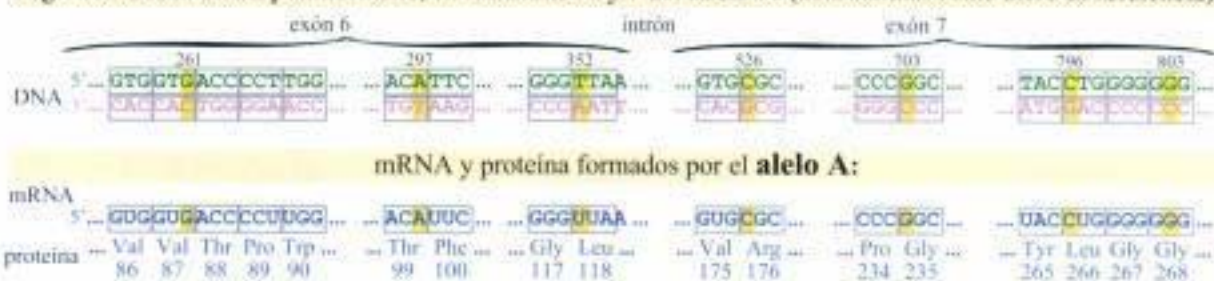
El gen H presenta dos alelos (H y h).

El producto génico del alelo dominante H es un polipéptido con actividad fucosiltransferasa.

El producto del alelo h no es funcional. (No se estudian aquí las diferencias entre H y h.)

Gen ABO

Para una mejor comprensión de las diferencias causantes del polimorfismo, se detallan los puntos polimórficos (en los exones 6 y 7) de las secuencias de mRNA y proteína sintetizados por cada alelo, considerando el A como referencia y el B y O como mutaciones de aquél.

Regiones causantes del polimorfismo, en los exones 6 y 7 del alelo A (considerado como alelo de referencia):

Los aspectos considerados se pueden resumir mostrando las relaciones entre los dos locus génicos, sus alelos, los productos génicos de éstos y los antígenos resultantes de la acción de esas enzimas:

Gen (o locus)	Alelo	Producto génico (proteína, glicosiltransferasa) (se indican varios nombres y su número de clasificación EC)	Antígeno (oligosacárido) generado por la glicosiltransferasa
Hh 2 alelos: H dominante, h recesivo	H	Fucosiltransferasa Galactósido 2-L-fucosiltransferasa EC 2.4.1.69	H
	h	No funcional	ninguno
ABO 3 alelos: A y B codominantes, O recesivo	A	Galactosaminiltransferasa Transferasa del grupo histo-sanguíneo A Glicoproteína-fucosilgalactósido α -N-acetilgalactosaminiltransferasa EC 2.4.1.40	A
	B	Galactosiltransferasa Transferasa del grupo histo-sanguíneo B Glicoproteína-fucosilgalactósido α -galactosiltransferasa EC 2.4.1.37	B
	O	No funcional	ninguno

El gen ABO constituye un ejemplo clásico de multialelismo: los alelos **A** y **B** son codominantes, mientras que **O** es recesivo, por lo que los 6 genotipos posibles determinan 4 fenotipos. En cuanto al gen H, posee sólo dos alelos, uno dominante (**H**) y el otro recesivo (**h**), que producen 3 genotipos y 2 fenotipos.

Grupos sanguíneos: genotipo, fenotipo, antígenos y anticuerpos presentes

Genotipo		Fenotipo	Glicosiltransferasas sintetizadas	Antígenos presentes	Anticuerpos en el plasma
HH o Hh	AA	grupo A	H y A	H (poco) y A	anti-B
	AO				
	BB	grupo B	H y B	H (poco) y B	anti-A
	BO				
	AB	grupo AB	H, A y B	H (poco), A y B	ninguno
	OO	grupo O	H	sólo H (mucho)	anti-A y anti-B
hh (*)	AA	grupo O	A	ninguno	anti-H, anti-A y anti-B
	AO				
	BB	grupo O	B	ninguno	anti-H, anti-A y anti-B
	BO				
	AB	grupo O	A y B	ninguno	anti-H, anti-A y anti-B
	OO	grupo O	ninguna	ninguno	anti-H, anti-A y anti-B

(*) En el caso de genotipo hh no se sintetiza el antígeno H por lo que no se pueden sintetizar los antígenos A ni B, aunque existan glicosiltransferasas de tipo A y/o B. Por tanto, el grupo sanguíneo es aparentemente O; se le llama fenotipo O_B Bombay.

26.4.2.2 Polimorfismo en el locus de la galactosemia

a) Descripción y bioquímica de la galactosemia

La galactosemia es una enfermedad autosómica recesiva causada por la imposibilidad de metabolizar la galactosa, que como consecuencia se acumula en sangre, síntoma que determina el nombre de la enfermedad. En teoría (véase la ruta metabólica), la galactosemia puede deberse a mutaciones en distintas enzimas; de hecho, así se ha demostrado, aunque con escasa incidencia. En la práctica, la causa más frecuente son mutaciones en el gen que codifica la enzima galactosa-1-fosfato uridililtransferasa (GALT). En situación normal, esta enzima, de 379 aminoácidos y una masa de 43 kDa, forma un dímero que cataliza la reacción



El nombre correcto de la enzima es uridililtransferasa, porque se transfiere a la galactosa-1-fosfato un resto uridilato (UMP); sin embargo, se suele usar uridiltransferasa. Obsérvese que ésta es sólo una de las 3 enzimas que transfieren un grupo uridilato.

Genotipo	Actividad enzimática	Fenotipo bioquímico
N N	100%	normal
G G	< 5%	galactosemia clásica
D D	50%	normal, "Duarte" (actividad reducida, sin transcendencia) (*)
LA LA	140%	normal, "Los Ángeles" (actividad excesiva, sin transcendencia) (*)
N G	50%	normal (actividad reducida, sin transcendencia)
N D	75%	normal (actividad reducida, sin transcendencia)
N LA	120%	normal (actividad excesiva, sin transcendencia)
G D	25%	situación límite
G LA	70%	normal (actividad reducida, sin transcendencia)
D LA	95%	normal

(*) El fenotipo bioquímico "Duarte" se identifica porque la enzima GALT se presenta en una isoforma más ácida (coherente con la mutación de asparragina a aspártico que se explica después), de punto isoeléctrico menor, que avanza más rápidamente en electroforesis y en isoelectroenfoque se sitúa más cerca del ánodo. Las variantes Duarte 1 y 2 se diferencian en su actividad enzimática.

Todos los fenotipos resumidos anteriormente son a su vez polimórficos. La reducción de la actividad enzimática varía en distinto grado según la combinación de alelos que presente el individuo.

Detalle de las mutaciones causantes del polimorfismo en el locus GALT

Se muestra la secuencia de la proteína GALT normal (locus génico en la banda cromosómica 9p13).

Los aminoácidos mutados que originan las variantes aparecen indicados en azul.

```

1  MSRSGTDPQQ RQQAEEADAA AATFRANDHQ HIRYNPLQDE WVLVSAHRMK RPWQQQVEPQ
61 LLKTVPRHDP LNPLCPGAIR ANGEVNPQYD STFLFDNDFP ALQPDAPSPG PSDHPLFQAK
121 SARGVCKVMC FHPWSDVTLF LMSVPEIRAV VDAWASVTEE LGAQYPWVQI FENKGAMMGC
181 SNPHPHCQVW ASSFLPDIAQ REERSQQAYK SQHGEPLLEME YSRQELLRKE RLVLTSSEHWL
241 VLVFPWATWP YQTLLEPRRH VRRLPELTPA ERDDLASIMK KLLTKYDNLF ETSFPYSMGW
301 HGAPTGSEAG ANWNHWQLHA HYYPLLRSA TVRKFMVGYE MLAQAQRDLT PEQAAERLRA
361 LPEVHYHLGQ KDRETATIA

```

La secuencia se indica mediante las abreviaturas de una letra para los aminoácidos (pág. 294).
Cada bloque contiene 10 aminoácidos (numeración del primero a la izquierda)

HPH (His-Pro-His) son aminoácidos del centro activo ("triada" característica; los dos residuos de histidina participan en la catálisis). Se ha identificado una mutación, **H184Q** (His→Gln, no descrita aquí), que afecta directamente al centro activo.

La mutación **Q188R** (Gln→Arg, debida a una transición CAG→CGG) es la más frecuente en individuos con galactosemia clásica, especialmente en la raza blanca europea y norteamericana. Conduce a una proteína con actividad enzimática casi nula (Gln-188 es un aminoácido altamente conservado, por tanto esencial para la función, probablemente por su cercanía al centro activo) y un fenotipo galactosémico.

La mutación **K285N** (Lys→Asn, debida a una transición AAG→AAT) es la segunda más frecuente en la población europea. Tiene también efectos fenotípicos graves (galactosemia).

La mutación **S135L** (Ser→Leu, debida a una transición TCG→TTG) es especialmente abundante en los casos de galactosemia entre la población norteamericana de raza negra (fenotipo "Negro"). La actividad enzimática se ve reducida en distinta medida según el tejido considerado. Se puede detectar este alelo mediante RFLP (pág. 381), aprovechando que la mutación hace desaparecer una diana para la enzima de restricción *TaqI*.

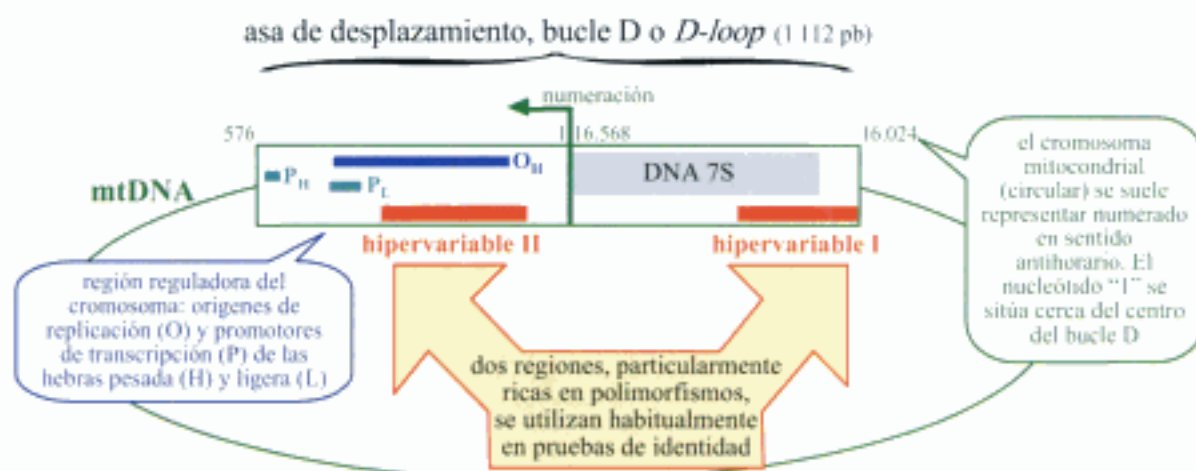
La mutación **N314D** (Asn→Asp, debida a una transición AAC→GAC) es característica de los fenotipos "Duarte" (o "Duarte-2") y "Los Ángeles" (o "Duarte-1"). El primero se caracteriza por una actividad enzimática reducida (cerca del 50%), mientras el segundo muestra actividad superior a la de individuos normales (140%). El fenotipo "Los Ángeles" suele presentar, además de la anterior, una mutación silenciosa **L218L** (Leu) por transición CTA→TTA, que parece ser responsable de una traducción más rápida y, por ello, de una mayor concentración de enzima que explicaría el aumento en actividad enzimática. Al igual que el caso anterior, se pueden detectar ambos alelos mediante RFLP, puesto que la mutación A→G hace aparecer una diana para la enzima de restricción *Ava II*.

más notables del genoma de un individuo es la exclusividad de su perfil genético (salvo para los hermanos gemelos univitelinos, individuos clónicos que tendrán los mismos alelos en cada uno de sus loci, aunque incluso en este caso pueden existir diferencias debidas a mutaciones somáticas). Se puede, por tanto, utilizar el análisis de estos polimorfismos para la identificación de una persona, con una fiabilidad igual o superior a la de las huellas dactilares, lo que ha dado lugar al término **huella genética** o **huella dactilar genética** (pág. 385). Con este objeto se emplean principalmente los polimorfismos en las regiones de DNA mini- y microsatélite (págs. 114 y 221). Por otra parte, puesto que las secuencias no codificantes no proporcionan información sobre características especiales de las personas (posibles enfermedades, cualidades físicas o psíquicas, etc.), el análisis basado en polimorfismos de regiones no codificantes posee la ventaja de evitar cualquier conflicto ético.

26.4.5 Polimorfismo en el genoma mitocondrial

En el genoma mitocondrial pueden darse igualmente mutaciones que generen un polimorfismo. Su DNA experimenta mayor proporción de mutaciones que el nuclear, debido a la exposición a especies reactivas de oxígeno, resultantes del metabolismo oxidativo mitocondrial, a encontrarse más desprotegido por carecer de histonas y a la menor eficacia de los sistemas de reparación mitocondriales. De esas mutaciones, aquellas que afectan a regiones codificantes (más del 93% de la molécula de mtDNA) suelen tener consecuencias perjudiciales para la funcionalidad de las proteínas que codifican y para la vida de la célula. Las mutaciones en la región no codificante se conservan con más facilidad, al no poseer efecto fenotípico. Aquellas mutaciones que no son perjudiciales dan lugar a polimorfismos genéticos; cada vez parece más claro que éstos tampoco son completamente inocuos para las diferentes funciones celulares. Por ejemplo, una combinación particular de alelos en varios loci polimórficos del cromosoma mitocondrial humano (poco frecuente y denominada "haplogrupo T") da lugar a espermatozoides con motilidad reducida.

Como consecuencia, los polimorfismos más comunes son aquéllos que afectan a la única región no codificante del DNA mitocondrial, conocida como **bucle D** (pág. 160). En este corto tramo de DNA (1.112 pb en humanos, menos del 7% del mtDNA total) se acumulan mutaciones con una frecuencia cinco veces superior a la del resto del DNA mitocondrial. En particular, la tasa de polimorfismos es máxima en dos pequeñas porciones del bucle D, denominadas por ello **regiones hipervariables**.



El análisis de polimorfismos en las regiones hipervariables mitocondriales está encontrando gran aplicación en estudios familiares, antropológicos evolutivos y también de identificación. A este éxito contribuye también la conveniencia de obtención y manejo de muestras mitocondriales, gracias a la existencia de un mayor número de copias (cientos de mitocondrias por célula, cada una con decenas de copias de mtDNA), lo que disminuye la probabilidad de que todas estén destruidas o dañadas, traduciéndose en definitiva en una mayor estabilidad del DNA. Por otro lado, al ser las regiones hipervariables mucho más cortas en el mtDNA es improbable su degradación en todas las copias de DNA de la muestra.

Tema 27

Análisis de genes. Detección y aplicaciones de los polimorfismos

27.1	Introducción	379
27.2	Análisis de genes	380
27.2.1	Búsqueda de genes por análisis de ligamiento	380
27.2.2	Empleo de métodos físicos para la identificación de genes	380
27.2.3	Identificación de genes por secuenciación	380
27.2.4	Clonación funcional	380
27.2.5	Identificación de genes y bioinformática	380
27.3	Detección del polimorfismo de secuencia de DNA	381
27.3.1	Polimorfismos en la longitud de los fragmentos de restricción (RFLP)	381
27.3.1.1	Ejemplo: Polimorfismo en la β -globina, causante de la anemia falciforme	382
27.3.2	Polimorfismo en el número de repeticiones en tándem (VNTR)	383
27.4	Aplicaciones de los polimorfismos de DNA	384
27.4.1	Mapa de polimorfismos SNP	384
27.4.2	Análisis del DNA con fines familiares	385
27.4.3	Huella dactilar de DNA	385
27.4.4	Análisis de polimorfismos con fines diagnósticos	386
27.4.4.1	Diagnóstico de errores congénitos del metabolismo	387
a)	Valoraciones preliminares	387
b)	Ensayos genéticos	387
c)	Incidencia de las enfermedades congénitas	387
27.4.5	Biochips o chips de DNA	387

27.1 INTRODUCCIÓN

Son muchos los aspectos de interés en el estudio del DNA génico, tanto desde el punto de vista básico (información sobre el propio gen) como, especialmente, para su aplicación al estudio de la enfermedad molecular con fines diagnósticos y terapéuticos. Aunque el conocimiento del factor genético no basta para resolver definitivamente la enfermedad, sí que es obligado conocer el aspecto molecular del gen o genes disfuncionales responsables (mutación o mutaciones) para poderla identificar. El número de enfermedades para las que ya se dispone de información génica es muy amplio; se pueden citar, por ejemplo, anemia falciforme, cáncer de mama y de colon hereditarios, corea de Huntington, distrofias musculares, enfermedad de Alzheimer, enfermedad de Tay-Sachs, fibrosis quística, neurofibromatosis, poliquistosis renal, síndrome de Marfan, síndrome del X frágil, sordera hereditaria, etc.

El objetivo de la genómica estructural (pág. 217) es identificar la secuencia del genoma completo y, sobre todo, localizar los genes responsables de enfermedades. Aunque el número de genes localizados ha crecido exponencialmente desde los años 70 hasta la actualidad, los ya identificados suponen menos del 10% del total de genes presentes en el genoma humano (estimado como media en unos 100.000 genes, pág. 239), que a su vez sólo equivale a un 3-5% del genoma total (pág. 108).

Entre los aspectos estudiados cabe citar la localización de genes en los cromosomas, la identificación de la función de sus productos génicos en la célula, la búsqueda y detección de genes anómalos responsables de enfermedades conocidas o nuevas, la identificación de mutaciones responsables de las anomalías, el estudio del polimorfismo de secuencia, el análisis de las proteínas sintetizadas en forma alterada, las pruebas genéticas en el

diagnóstico clínico de enfermedades monogénicas y multifactoriales, etc. El conocimiento de estos y otros aspectos se incrementará en los próximos años con los recientes avances en el conocimiento del genoma humano.

27.2 ANÁLISIS DE GENES

Bajo un término tan genérico como el de **análisis de genes** se estudian en este tema algunos de los planteamientos y herramientas básicas sobre las que se asienta la nueva medicina molecular, medicina genómica o medicina del futuro y en los que se basa la industria farmacéutica en un intento de desarrollar fármacos que actúen sobre genes anómalos, con estrategias para sustituirlos o corregir su función. Nuestro enfoque, inevitablemente restringido, no se dirige al estudio de enfermedades concretas, más propio de tratados especializados (clínicos, metodológicos, etc.), sino a la descripción de algunos aspectos aplicados en el diagnóstico y la terapéutica, partiendo de la información aportada en los temas precedentes y siguiendo como objetivo los puntos incluidos en los temas siguientes.

27.2.1 Búsqueda de genes por análisis de ligamiento

Un primer paso en la búsqueda de genes es la determinación aproximada de su presencia en los cromosomas. El análisis de ligamiento clásico, ya estudiado para el mapeo del genoma (pág. 218), permite averiguar la proximidad entre dos genes a partir de la frecuencia con que éstos aparecen recombinados en la descendencia. En humanos se consigue una localización más precisa si el análisis del gen buscado se hace no respecto a otro gen, sino a marcadores genéticos, especialmente los polimórficos (pág. 221).

27.2.2 Empleo de métodos físicos para la identificación de genes

Éstos proporcionan una localización más precisa de los genes, ya que se describen en distancias absolutas de pb (pág. 221). Se hace uso de métodos basados en algunas características de los genes, ausentes en regiones no codificantes.

- **Islotes CpG.** Se conoce la presencia de islotes CpG hipometilados (pág. 272) en la región promotora de genes de vertebrados que se transcriben activamente; más del 50% de los genes humanos presentan tales islotes. Es posible utilizar las genotecas (pág. 215) para buscar este tipo de secuencias y así identificar genes potenciales, aún no caracterizados.
- **Marcos abiertos de lectura.** Puesto que la región codificante de todo gen ha de empezar por un codón de inicio y finalizar en uno de terminación, la búsqueda en las genotecas de segmentos de DNA comprendidos entre estos tripletes, segmentos denominados **marcos abiertos de lectura** (ORF, de *open reading frame*; pág. 294), es otro método para encontrar genes potenciales.
- **Genotecas de cDNA.** Evidentemente, un gen que se expresa en un tipo celular determinado ha de estar representado por la presencia del mRNA transcrito a partir de él. El análisis de genotecas de cDNA (pág. 216) es, por tanto, otra fuente para la búsqueda de genes nuevos.

27.2.3 Identificación de genes por secuenciación

La secuenciación sería la forma más directa y sencilla de analizar un gen, de detectar un polimorfismo o una mutación. Sin embargo, en la práctica, este objetivo no es fácil dado el tamaño y complejidad del genoma humano y los problemas técnicos especializados e incluso económicos que comporta la secuenciación a gran escala. Es preferible seleccionar o "cribar", por alguno de los métodos anteriores, los fragmentos candidatos antes de secuenciarlos.

27.2.4 Clonación funcional

Se plantea cuando antes de conocer la alteración del gen se caracteriza su producto génico, por ejemplo la proteína responsable de una enfermedad (como la β^S -globina en la anemia falciforme). El objetivo es entonces identificar la secuencia del gen alterado a partir de la secuencia de la proteína anómala, proceso conocido como **clonación funcional**. A partir de la secuencia de aminoácidos se deduce la secuencia de nucleótidos del gen y se diseña una sonda (pág. 180) que permita localizar por hibridación el alelo mutante asociado a la enfermedad.

27.2.5 Identificación de genes y bioinformática

La Bioinformática o conjunto de tecnologías de la información aplicadas al estudio de los fenómenos biológicos está alcanzando su cenit con ocasión del Proyecto Genoma Humano. Ha sido espectacular la tecnología desarrollada para dilucidar el ordenamiento de los tramos de secuencia que se van descifrando. En este momento se habla ya de la bioinformática (o bioquimiinformática) como una ciencia que recoge información de áreas de conocimiento antes tan separadas entre sí como la bioquímica, la biología molecular, la estadística, la farmacología, la física, la fisiología, la

genética, la química física, etc., en colaboración con el desarrollo de las nuevas generaciones de ordenadores, internet, la robótica, etc.

A fin de identificar las secuencias que definen a los genes, se han elaborado los llamados **programas predictores de genes**; existen ya unos 20 de ellos, basados en distintas aproximaciones y combinando el empleo de sistemas de reglas, redes neuronales, inteligencia artificial, modelos estadísticos y, más recientemente, métodos lingüísticos.

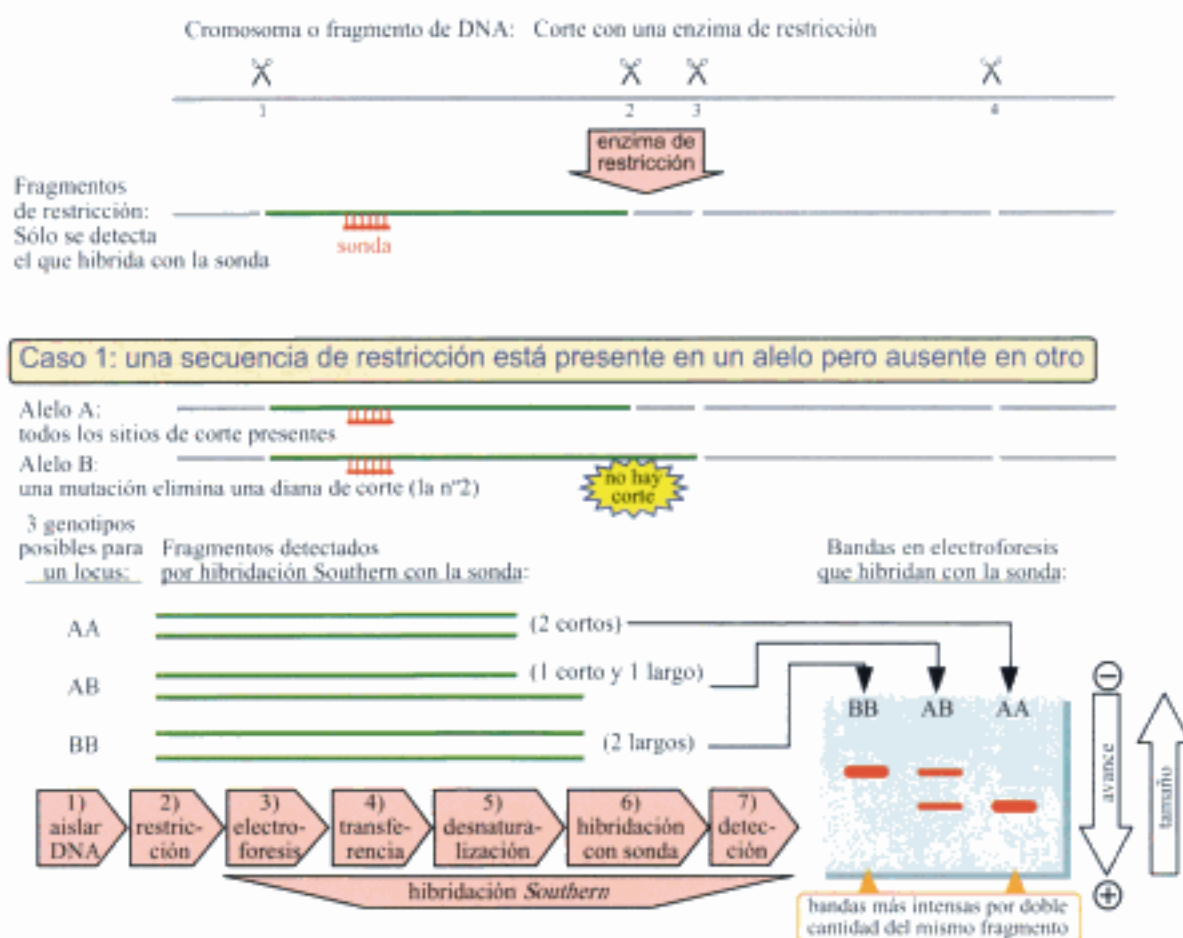
27.3 DETECCIÓN DEL POLIMORFISMO DE SECUENCIA DE DNA

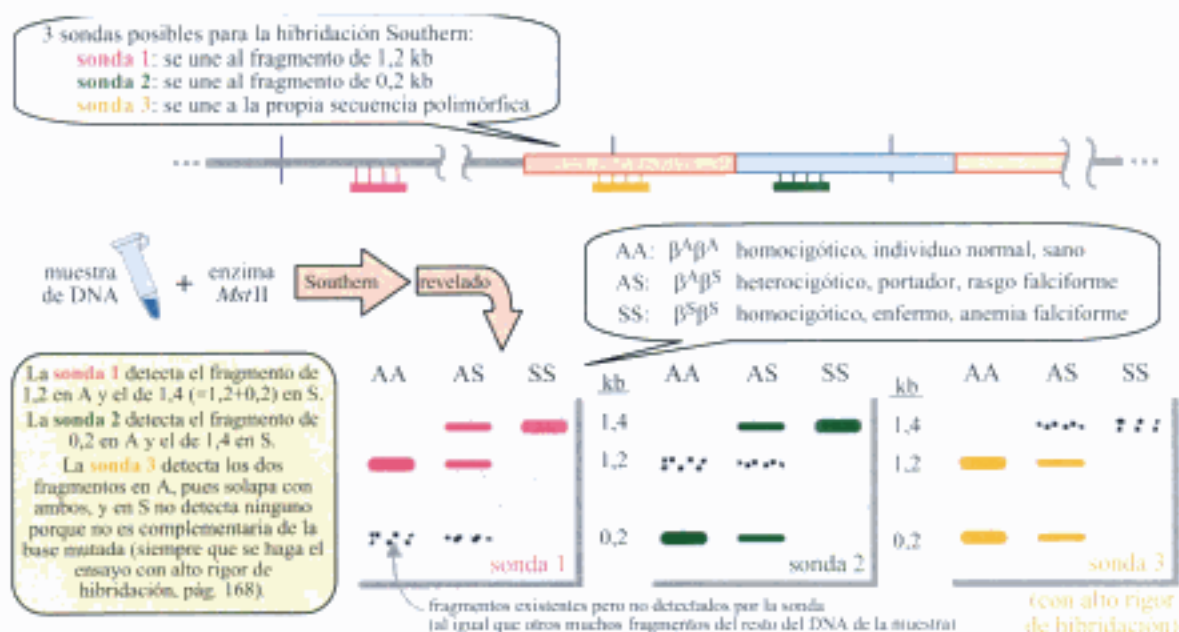
Para detectar las variaciones (polimorfismo, tema 26) en la secuencia del genoma de distintos individuos se acude a distintas estrategias con enfoques característicos, o una combinación de ellas. En principio, la forma más directa de detectar un polimorfismo es la secuenciación del DNA, pero por su laboriosidad se suele acudir a otras técnicas, entre las que destacan los análisis RFLP y VNTR.

27.3.1 Polimorfismos en la longitud de los fragmentos de restricción (RFLP)

Quizá el método más empleado, se denomina simplemente PLFR o **RFLP** (*Restriction Fragment Length Polymorphism*). No se basa en un cambio fenotípico, un cambio identificable de la expresión de caracteres morfológicos o funcionales, sino en la detección de aquellas variaciones de la secuencia del DNA (codificante o no) que tienen como consecuencia un cambio en una diana de restricción. Por tanto, los fragmentos de restricción que se obtienen son diferentes dependiendo del alelo presente.

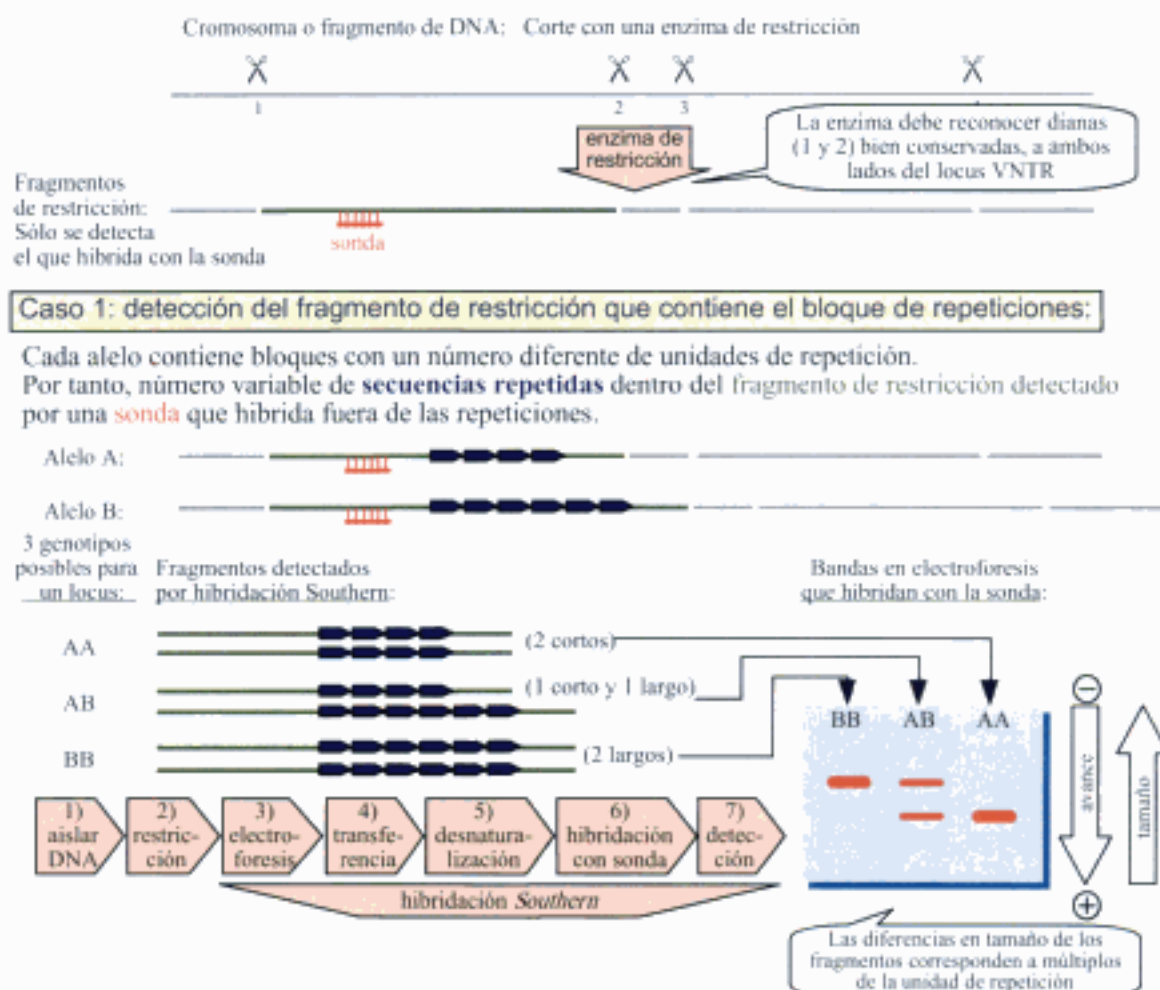
Cada muestra de DNA generará múltiples fragmentos, de longitud diferente en función de los sitios de restricción presentes. Puesto que muchos de ellos no corresponden a la región que se pretende estudiar, generalmente se utiliza una sonda que hibrida en posición cercana al sitio polimórfico y se analizan los fragmentos por el método de Southern (pág. 171). Se puede por ello afirmar que el RFLP es un análisis de polimorfismos basado en la hibridación.





27.3.2 Polimorfismo en el número de repeticiones en tándem (VNTR)

Este tipo de polimorfismo surge de la presencia de DNA repetitivo no codificante en eucariotas, concretamente de repeticiones en tándem (DNA satélite, minisatélite y microsatélite, págs. 113-114). El polimorfismo, entre individuos y entre los dos alelos de un mismo individuo, surge porque el número de unidades de repetición que forman cada bloque no es siempre el mismo; de ahí el nombre, polimorfismo de número variable de repeticiones en tándem, o simplemente NVRT o **VNTR** (*Variable Number of Tandem Repeats*). La razón de que varíe con tanta facilidad el número de repeticiones está en la posibilidad de recombinación entre ellas con sobrecruzamientos desiguales (pág. 367).



Caso 2: detección del propio bloque de repeticiones:

En este caso, la **sonda** hibrida con la secuencia repetida. El resultado es similar al del caso 1.



Los polimorfismos VNTR que se emplean en la práctica corresponden a bloques moderadamente grandes de unidades de 5 a 64 pb, correspondientes a los **minisatélites hipervariables**. Otros tipos no suelen recibir ese nombre, como el polimorfismo de repeticiones sencillas en tándem (unidad de repetición de 1 a 4 pb, es decir, DNA microsatélite) y el polimorfismo de repeticiones en tándem asociado con grandes bloques de DNA satélite. Los **loci hipervariables** se llaman así por tener de forma característica muchos alelos. Los marcadores más informativos tienen decenas de alelos, por lo que es poco probable que dos individuos no emparentados compartan los mismos. Estos marcadores VNTR proporcionan mucha información para el análisis de ligamiento genético (mapas genéticos, pág. 218) y para la identificación de individuos (pruebas forenses y de paternidad, pág. 385).

Tanto los RFLP como los VNTR han incrementado fuertemente el número de marcadores polimórficos de que se dispone para examinar la diversidad genética entre individuos y para seguir la segregación de diferentes porciones del genoma a lo largo de las familias. Se han descrito miles de estos polimorfismos, localizados en todos los cromosomas.

Como alternativa al método de Southern, para detectar el polimorfismo VNTR puede aplicarse una amplificación por PCR (pág. 188). En lugar de una sonda, se precisan entonces dos cebadores que hibriden a sendos lados del bloque de repeticiones. Puesto que así sólo se amplifica la región que contiene el locus polimórfico, no es necesario purificar el DNA, ni digerirlo con enzimas de restricción, y en la electroforesis sólo se observarán (sin hibridación, basta la tinción normal de DNA, pág. 138) las bandas de los fragmentos que incluyen el bloque de repeticiones, de tamaños distintos según el alelo presente.

27.4 APLICACIONES DE LOS POLIMORFISMOS DE DNA

El análisis del polimorfismo genético da pie a numerosas aplicaciones, tanto en investigación básica como en clínica, que se basan en los tipos de polimorfismos antes descritos.

- Los polimorfismos del DNA que **no tienen efecto fenotípico** se utilizan para la comprobación de la identidad (huella dactilar genética). Se emplean para ello principalmente los RFLP y VNTR. Por ejemplo:
 - ♦ Diagnóstico de la paternidad biológica y seguimiento de árboles genealógicos.
 - ♦ Identificación de sospechosos en procedimientos penales, por comparación con el DNA procedente de diversos restos biológicos (sangre, semen, saliva, raíces de cabello, tejidos corporales diversos, piezas dentales, etc.).
 - ♦ Identificación post-mórtem de individuos, en casos judiciales o catástrofes.
 - ♦ Otras aplicaciones de interés diverso. Por ejemplo, es posible aplicar estas metodologías a muestras clínicas con el fin de comprobar identidad de las biopsias, posibilidades de trasplante, inestabilidad genética característica de ciertos tumores, etc.
- El estudio de los polimorfismos que tienen una contribución mínima o leve a la **susceptibilidad** frente a ciertos procesos patológicos posee interés médico, especialmente en la prevención de enfermedades complejas.
- Los polimorfismos que desempeñan un **papel directo** en la aparición de un fenotipo patológico son, evidentemente, aplicables para el diagnóstico, así como para el estudio de los mecanismos moleculares de la enfermedad (Patología Molecular).

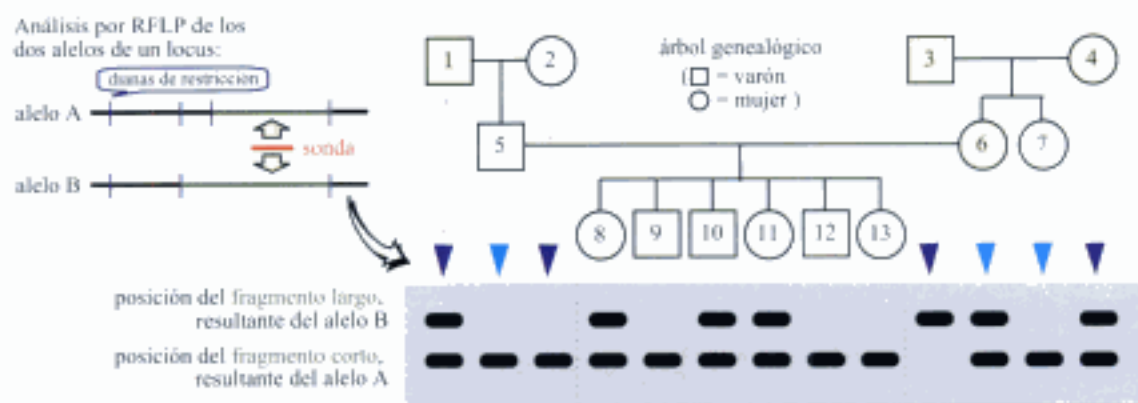
27.4.1 Mapa de polimorfismos SNP

Los polimorfismos SNP o de un solo nucleótido (pág. 365) son responsables de una gran parte de la diversidad del genoma humano. Aunque la mayoría de ellos no originan directamente enfermedades, en ocasiones pueden localizarse muy cerca de mutaciones o polimorfismos involucrados en procesos patógenos, lo que los hace útiles como marcadores genéticos (pág. 221). Al irse caracterizando cada vez un número mayor de estos polimorfismos, se hace posible la construcción de un **mapa de SNPs**, lo que se une a la utilidad propia de algunos como determinantes de enfermedades.

Se ha iniciado un gran proyecto para identificar los SNPs del genoma humano y lograr un mapa de alta resolución con la posición de unos 170.000 de ellos. En él participan, desde abril de 1999, tres centros de investigación, avalados por 10 grandes empresas farmacéuticas, el Wellcome Trust británico y Motorola. En abril de 2000, con una inversión que supera los 5.000 millones de pesetas, se habían identificado y cartografiado más de 41.000 SNPs, a partir del material genético de 24 individuos de distintas etnias. En este proyecto se ha acordado no patentar las secuencias descubiertas.

27.4.2 Análisis del DNA con fines familiares

El análisis, generalmente por RFLP y VNTR, de unos cuantos polimorfismos permite realizar un seguimiento de la herencia de los alelos respectivos y así obtener conclusiones sobre las relaciones familiares entre varias personas. Análogamente, se pueden realizar estudios evolutivos y antropológicos, tales como el seguimiento de las migraciones humanas desde tiempos prehistóricos y las relaciones entre los antecesores de la raza humana.



Análisis:

- 7 no es hija de 3 (le falta el alelo B paterno).

Si B está asociado a una enfermedad autosómica recesiva (símbolos coloreados):

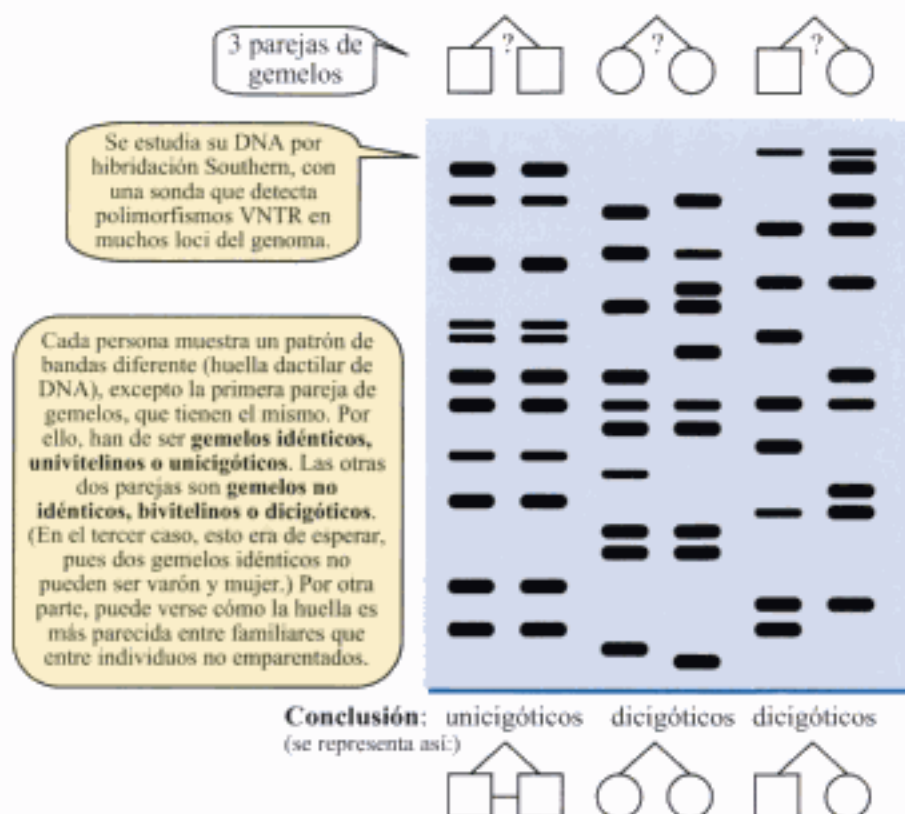
- 3 la padeció (homocigótico), mientras que 1, 4 y 6 eran portadores (heterocigóticos).
- 5 heredó de su padre (1) el alelo A, con lo que se extinguió la herencia de la enfermedad por esa rama.
- El gen portador de la enfermedad de 6 fue heredado de su padre (3), no de la madre (4).
- De la última generación, 8, 10 y 11 son portadores, mientras que 9, 12 y 13 están sanos.

Una de las principales aplicaciones de estos estudios familiares ha sido relacionar la presencia de una enfermedad genética, causada por un defecto en un gen aún desconocido, con la de un polimorfismo determinado. Mediante un estudio amplio del árbol familiar, basado en la transmisión mendeliana del gen patológico y de los marcadores polimórficos, se puede llegar a encontrar un alelo polimórfico que esté ligado a la enfermedad (por la proximidad de sus loci respectivos en el genoma, pág. 218). Esto permite, en primer lugar, identificar de forma aproximada la posición del gen patológico y así facilitar su búsqueda y, por otro lado, el diagnóstico preventivo de la enfermedad mediante la detección del alelo polimórfico asociado. Este procedimiento se ha aplicado con algunas enfermedades humanas, como la fibrosis quística y la corea de Huntington.

27.4.3 Huella dactilar de DNA

Como se ha visto hasta ahora, los polimorfismos RFLP y VNTR se localizan con una sonda particular, para examinar un locus único. Sin embargo, la mayoría de secuencias repetitivas que se estudian en los polimorfismos VNTR aparecen en múltiples zonas del genoma (recuérdese que los mini- y microsatélites son DNA moderadamente repetitivo y disperso, organizado en bloques dispersos de repeticiones en tándem, pág. 114). Las unidades de repetición presentes en puntos distintos del genoma pueden ser idénticas o bien lo suficientemente similares como para permitir su detección simultánea en todos esos loci. Si se utiliza como sonda una secuencia que pueda hibridar con todos los loci de una familia de DNA mini- o microsatélite, se puede observar simultáneamente el patrón polimórfico de todos ellos que, dada la combinación de gran número de variaciones individuales, es prácticamente único para cada individuo. Por ello, al patrón resultante se le llama **"huella dactilar de DNA"** (*DNA fingerprint*). Sólo los gemelos idénticos presentan huellas indistinguibles.

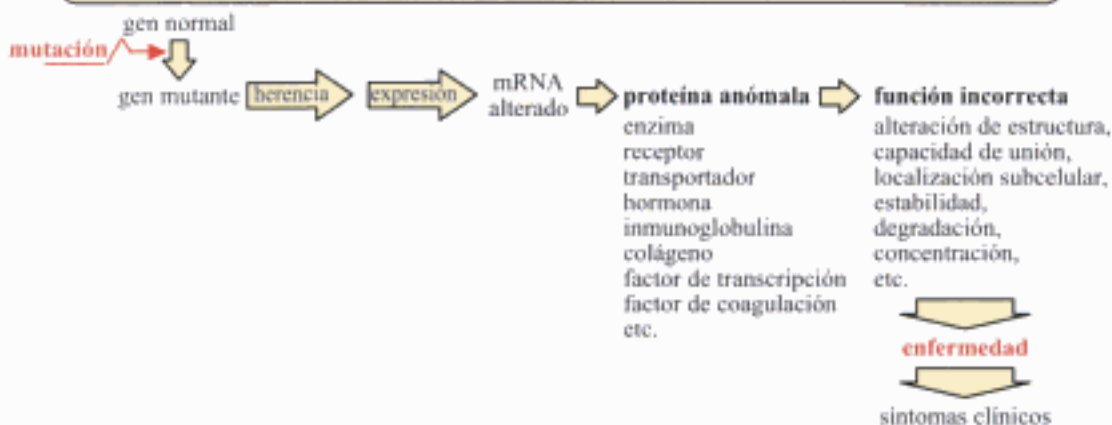
La huella genética puede utilizarse en estudios familiares, pruebas de paternidad, identificación de recién nacidos, criminología, pruebas forenses, etc. Otra de sus aplicaciones es la detección rápida de cambios somáticos del genoma, por ejemplo, al comparar DNA de tejidos normal y neoplásico. En una alta proporción de tumores malignos, la huella genética sufre cambios, que consisten en la aparición o pérdida de bandas.



27.4.4 Análisis de polimorfismos con fines diagnósticos

El análisis de polimorfismos es un método óptimo para el diagnóstico, en especial por la posibilidad de detectar la enfermedad antes de que se desarrolle, o la predisposición genética a ciertas enfermedades cuya causa no es única. Es de enorme interés, por ejemplo, el diagnóstico prenatal de enfermedades congénitas, es decir de "toda anomalía del desarrollo morfológico, estructural, funcional o molecular presente al nacer (aunque pueda manifestarse más tarde), externa o interna, familiar o esporádica, hereditaria o no, única o múltiple" (definición de la OMS).

Esquema general de la secuencia de acontecimientos de la enfermedad genética.
Por ejemplo, para una enfermedad monogénica (patrón de herencia mendeliana por alteración en un punto único del gen, que afecta a una sola proteína).



Una vez realizada la exploración médica del paciente, es decir delineado el fenotipo clínico, la contribución del laboratorio al diagnóstico puede hacerse a tres niveles: expresión y herencia de los genes (nivel génico o genoma), efecto sobre el producto génico anormal resultante (lo que hoy se llama proteoma), y vía metabólica o desequilibrio químico afectado (nivel bioquímico). Cualquiera de estas situaciones constituye un objetivo y debe ser estudiado en Patología Molecular, una materia hoy día multidisciplinar.

Los errores congénitos se pueden considerar como paradigma de los trastornos bioquímicos determinados genéticamente en la estructura y función de moléculas proteicas. La enfermedad es simplemente la ilustración más evidente de una variación genética que afecta a la salud.

27.4.4.1 Diagnóstico de errores congénitos del metabolismo

A diferencia de adultos, donde se suelen utilizar muestras de sangre u otros tejidos, en el diagnóstico prenatal se utilizan dos tipos de técnicas: no invasivas, que incluyen ultrasonidos y radiografía, e invasivas, como amniocentesis, biopsia de vellosidades coriónicas, cordocentesis o biopsias de hígado fetal (págs. 119-120). Si el contenido en DNA de la muestra no es suficiente, puede amplificarse la región de interés por PCR (tema 15).

a) Valoraciones preliminares

Entre los parámetros de carácter básico o preliminar para la monitorización fetal y el diagnóstico prenatal se emplean la **gonadotropina coriónica** (hCG), una glicoproteína producida por el embrión en desarrollo, que se detecta como indicador de embarazo desde los primeros días de la concepción, el **lactógeno placentario** (hPL) sintetizado por la placenta, que aumenta hasta el término del embarazo en sangre materna, y el **estriol**, un esteroide sintetizado por feto y placenta. Hoy día, estos ensayos han sido superados por la investigación ultrasónica y cardiotocográfica. Para establecer riesgos de malformación fetal se mide la **α -fetoproteína** (AFP), sintetizada en saco vitelino e hígado fetal, que debido a su pequeño tamaño pasa del plasma a la orina fetal y, por tanto, al fluido amniótico y sangre materna. Un aumento por encima del valor normal es índice de anomalías congénitas, especialmente defectos del cierre del tubo neural (espina bífida, anencefalia, etc.). Además de estos parámetros, pueden también medirse, con distintos fines diagnósticos, los niveles de bilirrubina y concentración de fosfolípidos en fluido amniótico, el pH, gases sanguíneos y concentración de lactato en sangre fetal, etc.

b) Ensayos genéticos

El ensayo genético detectará la patología buscando, como ya se ha descrito, el alelo asociado a la enfermedad, si se conoce su gen, o mediante otro polimorfismo cuyo locus muestre ligamiento con el locus patológico desconocido. Se aplican los métodos ya descritos, basados en el ligamiento de genes, polimorfismos RFLP y VNTR, reacción de la PCR, etc. Como ejemplos cabe citar:

- Medida de un sitio de restricción alterado por la mutación, como en el polimorfismo RFLP con Mst II en la anemia falciforme (pág. 382).
- Empleo de oligos sintéticos como sondas para la secuencia alterada, cuando el trastorno genético se debe a un cambio de un nucleótido específico. Un ejemplo es la deficiencia en α_1 -antitripsina, que conduce a una mayor probabilidad de desarrollar enfisema pulmonar, debido a un cambio en una sola base (pág. 376). Un análisis Southern con una sonda sintética permite determinar si el DNA contiene la secuencia normal o la mutada.
- Ensayos por PCR. Puesto que esta técnica permite amplificar la secuencia de DNA potencialmente defectuosa, puede emplearse para el diagnóstico prenatal de enfermedades donde se encuentra en cuestión el sitio exacto de mutación.

c) Incidencia de las enfermedades congénitas

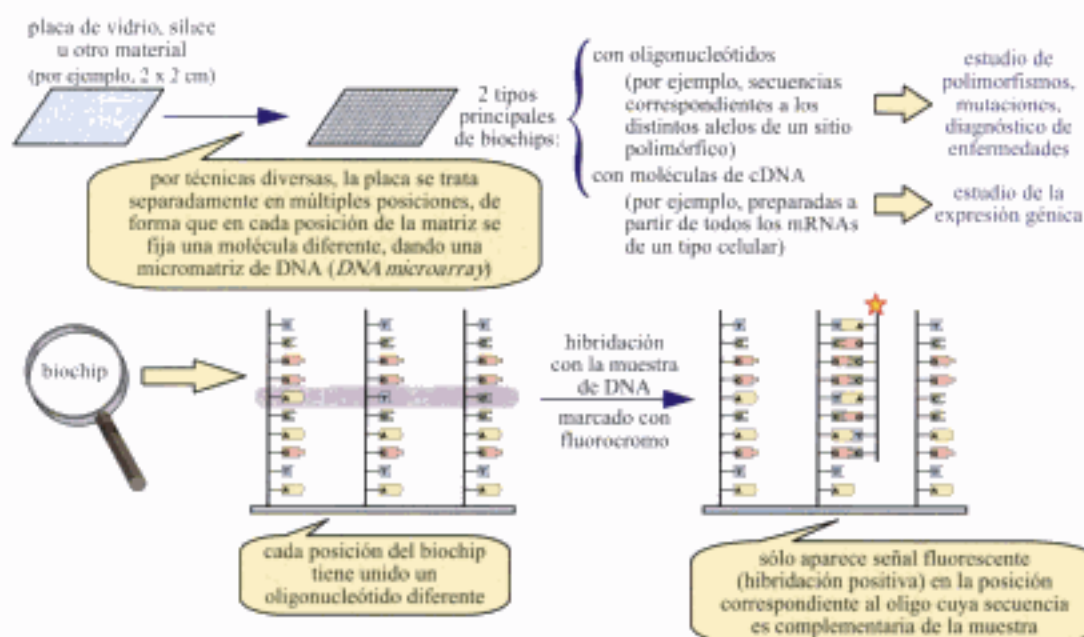
Se han descrito cerca de 4.000 fenotipos monogénicos, de los cuales la mitad constituyen trastornos genéticos. En centenares de estas enfermedades, tanto las comunes como las más raras, se ha reconocido el defecto bioquímico básico y, en muchos casos, se ha aislado y clonado el gen responsable. Esto quiere decir que existe hoy información de algunos o varios de los tres niveles antes descritos, para unos 500 trastornos con herencia mendeliana. Su incidencia a nivel mundial es de, aproximadamente, uno de cada 500-800 niños.

Es precisamente esta elevada incidencia de errores congénitos del metabolismo lo que ha motivado la aparición de **programas de cribado (screening) neonatal** en todos los países desarrollados. En España se realizan ya en todas las Comunidades Autónomas. Su objetivo es diagnosticar lo más precozmente posible la alteración metabólica, antes de que se manifieste en lesiones irreversibles. Esta detección supone iniciar un tratamiento precoz y más eficaz, tratando de favorecer el desarrollo de una vida normal. Incluso en casos, por ejemplo la fibrosis quística, en que la enfermedad no se puede evitar, no sólo se puede conocer precozmente sino que se pueden prevenir las infecciones o tratar la insuficiencia pancreática.

27.4.5 Biochips o chips de DNA

El conocimiento de la causa de las enfermedades genéticas en familias se ha acelerado gracias al análisis de marcadores genéticos. En el futuro se podrá disponer de marcadores más fiables y sofisticados para diagnosticar precozmente la alteración molecular, a nivel celular, antes de que deje de funcionar el órgano. El diagnóstico evoluciona hacia el conocimiento de un catálogo de variaciones genéticas mediante una aplicación tecnificada, fruto de estos y otros avances en conjunción con la bioinformática. Se trata de buscar un perfil genético que nos permita identificar las enfermedades a las que podemos estar predispuestos (por ejemplo, tumor maligno, patología cardiovascular, hepatitis, enfermedad neurodegenerativa), desde el punto de vista de las mutaciones. En definitiva, va a suponer un gran paso hacia la **medicina predictiva**, que permitirá tomar medidas terapéuticas precoces para eliminar el riesgo de enfermedad.

Los llamados **biochips**, **genochips** o **genosensores**, fruto de la fusión de la biología molecular con la automatización y la informática, están destinados a revolucionar el diagnóstico molecular del futuro, al discriminar de una manera muy sencilla, rápida y automatizada la presencia de la enfermedad.



Se trata de placas de silicio, vidrio u otro material, de unos centímetros de tamaño, a cuya superficie se han adherido según una disposición ordenada, conocida, bien diversas moléculas de cDNA o bien oligonucleótidos de secuencias ligeramente diferentes, específicas para un determinado gen o una región de DNA de interés. La tecnología empleada permite unir en cada biochip entre cientos y decenas de miles de moléculas diferentes, por lo que se puede hacer un número de ensayos muy elevado de forma simultánea a partir de una muestra muy pequeña (por ejemplo, una gota de sangre o menos de un microgramo de mRNA total purificado de un tipo celular concreto). Esencialmente, se trata de un ensayo de hibridación (tema 13), en el que el papel de sonda lo desempeñan las moléculas unidas al biochip y en este caso se marca el DNA o RNA de la muestra. La muestra, desnaturalizada, se vierte sobre el biochip para que hibride con las secuencias de éste, se lava y se detecta en qué posiciones aparece el marcaje.

Una vez superado el afán por el estudio exclusivo de la secuencia del genoma humano y admitido que ésta no es la principal cuestión, se ha entrado ya en la necesidad de aplicar los nuevos conocimientos a la identificación de los genes contenidos en dicha secuencia y sus alteraciones, es decir, las mutaciones y su detección con fines diagnósticos. Posteriormente, habrá de darse explicación a la función de los genes, lo que algunos llaman "darle sentido al genoma", analizando la expresión de los mRNAs que una célula produce en un momento dado ("transcriptoma"). En estos dos áreas (mutaciones y expresión génica) se basa el progreso del diagnóstico del futuro, y es precisamente en ambos campos donde adquiere interés la aplicación de los biochips, inicialmente desarrollados para identificar secuencias de DNA. Se ha conseguido ya preparar biochips de tamaño muy pequeño (laminillas de vidrio del tamaño de una uña), donde se deposita una fina capa de los cDNAs obtenidos a partir de los mRNAs de un tipo de célula, etiquetados adecuadamente. Una vez añadida la muestra con el mRNA problema, se observa con cuáles de los cDNAs del biochip hibrida. Con estos biochips se pretende identificar miles de mRNAs humanos relacionados con enfermedades.

Tema 28

Enfermedades moleculares: enfermedades monogénicas

28.1	Concepto y clasificación	389
28.2	Enfermedades exógenas	390
28.3	Enfermedades genéticas	390
28.3.1	Extensión y frecuencia de las enfermedades genéticas	391
28.3.2	Clasificación de las enfermedades genéticas	391
28.3.2.1	Tipo de célula que sufre la alteración	391
28.3.2.2	Producto génico defectuoso que da lugar al trastorno	392
28.3.2.3	Extensión de la alteración	392
28.3.2.4	Tipo de cromosoma afectado	392
28.3.3	Enfermedades monogénicas nucleares	393
28.3.3.1	Ejemplos de enfermedades monogénicas	393
28.3.3.2	Transmisión en familias	393
28.3.3.3	Clasificación de los trastornos monogénicos	394
a)	Enfermedades autosómicas dominantes	395
b)	Enfermedades autosómicas recesivas	395
c)	Enfermedades dominantes ligadas al cromosoma X	396
d)	Enfermedades recesivas ligadas al cromosoma X	397
e)	Enfermedades ligadas al cromosoma Y	397
28.3.4	Enfermedades mitocondriales	398
28.4	Enfermedades poligénicas, multifactoriales, mixtas o complejas	399

28.1 CONCEPTO Y CLASIFICACIÓN

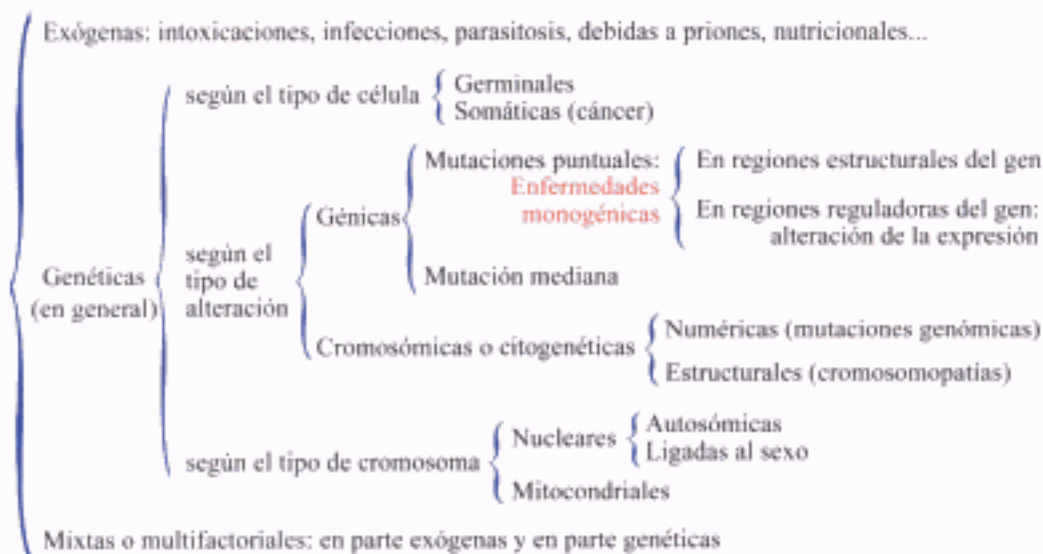
Se admite hoy día que prácticamente todas las enfermedades tienen una **causa molecular**. Cualquier enfermedad, incluyendo las conocidas tradicionalmente, cursa en mayor o menor medida con alteraciones en la estructura, propiedades, metabolismo o función de las biomoléculas. Entre las causas que inducen la enfermedad se encuentran agentes físicos, ambientales o no; agentes biológicos como virus, bacterias, parásitos, etc., y trastornos genéticos, hormonales, inmunológicos, nutricionales, etc. Todos ellos se traducen en disfunciones tisulares y celulares, temporales o permanentes, cambios en el medio interno, variaciones en la función de órganos y sistemas y se reflejan, en último término, en forma de enfermedad. Incluso aquellas debidas a agentes externos (traumatismos por accidentes, catástrofes, etc.), que obviamente son consideradas como de causa no molecular, se convierten una vez iniciadas en enfermedades moleculares. De ahí que un estudio moderno de la enfermedad requiera la consideración de los aspectos moleculares que acompañan a las observaciones clínicas y se consideren válidos, complementarios y no excluyentes los dos tipos de clasificación de la enfermedad:

- Las **clasificaciones clínicas** se basan en el diagnóstico a partir de la observación médica (exploración) con el apoyo de los datos de laboratorio (magnitudes biológicas, isótopos, radiología, etc.). Desde el punto de vista genético, se basan en un criterio fenotípico, la manifestación bajo la forma de síndrome clínico. Este tipo de estudio es exclusivo de los profesionales clínicos.
- Las **clasificaciones moleculares**, o estudio de la enfermedad bajo un enfoque molecular, se basan exclusivamente en criterios genotípicos. Dependiendo de las causas que alteran la constitución del genoma se distinguen tres tipos: exógenas, genéticas o mixtas. Al no conocerse todavía con precisión la causa molecular de muchas enfermedades,

estas clasificaciones son aún imperfectas. Este tipo de estudio lo pueden realizar hoy día diversos profesionales especializados en el área de la biomedicina.

El objetivo de la **patología molecular** es precisamente el conocimiento de la enfermedad desde el punto de vista de su alteración molecular, para contribuir al diagnóstico y la terapia. Empieza a observarse a comienzos del siglo XX, al establecerse el término **errores congénitos del metabolismo** para describir aquellas alteraciones hereditarias, perennes durante toda la vida del paciente, que afectan a rutas específicas del metabolismo por deberse a la ausencia o inactivación de enzimas específicas (entre otras, albinismo, alcaptonuria, cistinuria y galactosemia). El inicio real y sistemático de la patología molecular, que tiene lugar en los años 40-50, se desarrolla con las técnicas de estudio de los cromosomas humanos y su papel en el desarrollo sexual y en las anomalías cromosómicas relacionadas; en España, la patología molecular se introduce en la clínica en los años 60. El rastreo de determinadas enfermedades en recién nacidos y la ayuda de nuevas técnicas de manipulación y análisis del DNA, a partir de los 70, transformó la genética médica. Al final de los años 80 se habían recopilado ya varios miles de alteraciones genéticas hereditarias distintas, casi siempre con aumento o disminución de la actividad de una enzima específica. Bajo el término **genómica** se incluye hoy día no sólo el estudio del genoma, sino también la aplicación de exámenes genéticos para identificar las alteraciones responsables de una enfermedad, detectar en una población las personas con riesgo y así contribuir al tratamiento de la enfermedad. De forma análoga, en el futuro la **proteómica** reafirmará el papel que desempeñaron las proteínas, en solitario, en el inicio de la patología molecular.

Clasificación molecular de las enfermedades humanas



28.2 ENFERMEDADES EXÓGENAS

Se denominan así las enfermedades causadas por agentes de todo tipo, externos al individuo: intoxicaciones por productos químicos, drogas o medicamentos, parásitos unicelulares o pluricelulares, toxinas biológicas, agentes infecciosos como bacterias y virus, etc. También se incluyen las enfermedades por alteraciones nutritivas (generalmente carenciales) y las causadas por priones (proteínas de plegamiento anómalo capaces de inducir el mismo en otras moléculas de proteína normal y por ello consideradas "proteínas infecciosas"). Ninguna de ellas se trata de forma específica en este libro.

28.3 ENFERMEDADES GENÉTICAS

Son el grupo de enfermedades moleculares por excelencia. Para su inclusión en este grupo, la enfermedad ha de cumplir las siguientes características:

- Debe estar determinada genéticamente, es decir, deberse a mutaciones que alteran la secuencia o la organización del genoma codificante.
- Debe afectar a moléculas en cantidad, estructura, actividad o función. En especial, se alteran las proteínas, en cantidad o en actividad.
- La alteración bioquímica causante de la enfermedad debe ser el resultado de las alteraciones en las moléculas, que afectan a estructuras celulares o vías metabólicas en las que se encuentran implicadas (en especial, las proteínas de carácter enzimático).

28.3.1 Extensión y frecuencia de las enfermedades genéticas

En cuanto a la **extensión** de la enfermedad, debe diferenciarse entre individuos **portadores** de defectos genéticos, muy abundantes, pero en los que no se hace evidente la enfermedad por la existencia de un alelo sano que suplente la función del gen alterado, y **enfermos** o afectados por el trastorno genético. Este último caso corresponde a las enfermedades genéticas propiamente dichas, de gran importancia globalmente a pesar de la baja frecuencia de cada una. Los datos suelen ser inexactos, entre otros motivos por errores diagnósticos y por la complejidad del estudio, especialmente en las enfermedades multifactoriales.

Algunas cifras relativas a las enfermedades genéticas en general

- Están influidas por genes un 30% de todas las enfermedades humanas.
- Se calculan unas 7.000 enfermedades con base genética.
- En muchos de los 4.500 loci descritos se conoce al menos una forma de mutación asociada a un trastorno clínico significativo.
- Hay defectos de genes aislados en un 6-8% de las enfermedades infantiles.
- Un 5% de las personas menores de 25 años presenta un trastorno genético.
- Son también la causa de muchas enfermedades crónicas de adultos.
- Si se incluyen los rasgos complejos de aparición tardía, un 60% de personas presenta enfermedades con influencia genética en algún momento de su vida.

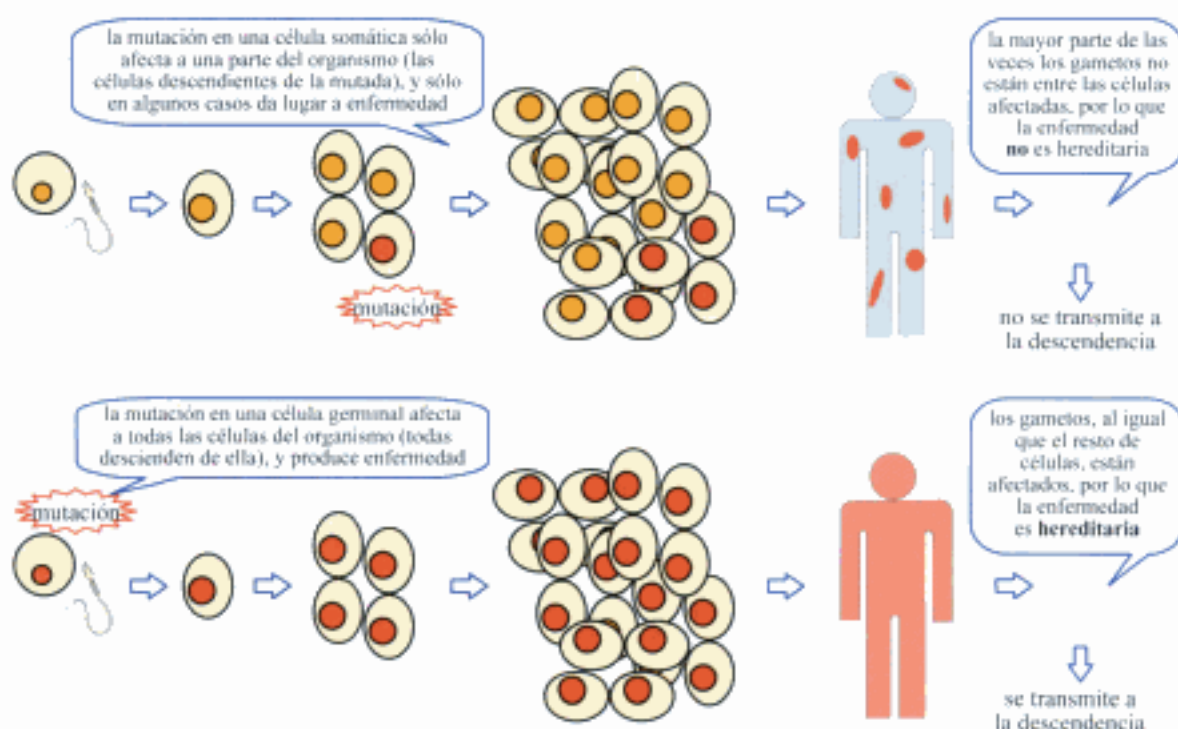
La frecuencia con la que aparece una enfermedad se expresa mediante dos magnitudes epidemiológicas: incidencia y prevalencia. La **incidencia** es la fracción del número de casos respecto al número de individuos nacidos vivos; puede resultar difícil de estimar si la enfermedad no se observa físicamente o no se diagnostica en el momento del nacimiento. La **prevalencia**, o frecuencia de casos en una población de individuos vivos, es aún más difícil de estimar, sobre todo cuando la población es grande y la enfermedad rara.

28.3.2 Clasificación de las enfermedades genéticas

Se pueden seguir varios criterios paralelos:

28.3.2.1 Tipo de célula que sufre la alteración

Las enfermedades que se transmiten de generación en generación son de carácter hereditario debido a que se originan por una alteración en las **células germinales** (pág. 344). La alteración de las **células somáticas** no es hereditaria ni afecta a todo el organismo, y da lugar a pocas enfermedades. Entre éstas debe citarse el cáncer (tema 30), que se estudiará como ejemplo de enfermedad somática, a la vez que de carácter mixto o multifactorial.



28.3.2.2 Producto génico defectuoso que da lugar al trastorno

Este criterio ha permitido estudiar muchas enfermedades genéticas y establecer las clasificaciones propias de la patología molecular tradicional, con muchas variantes según el tipo de proteína alterada.

- La enfermedad se denomina **disenzimía** cuando se altera una enzima. Constituye el caso más frecuente de estudio de una vía metabólica alterada. Se modifican sus características cinéticas o su regulación (sensibilidad a activadores e inhibidores). Si disminuye la actividad, puede desaparecer el producto de la reacción catalizada y, como consecuencia, el producto final de la vía metabólica a la que pertenece la enzima, y también puede acumularse el sustrato correspondiente o un precursor.
- La anomalía puede afectar a una proteína implicada en procesos de **transporte** de moléculas por el organismo (citósol, sangre o medio extracelular) o a través de membranas (plasmática o intracelulares).
- Alteraciones de la **biosíntesis** o la **degradación** de otras proteínas, a través de mecanismos estructurales o procesos metabólicos, tales como transcripción, procesamiento postranscripcional, traducción, plegamiento, modificación postraduccional, agrupación de subunidades, tráfico, formación de complejos multiproteicos, etc.
- Alteración de proteínas **estructurales**, por ejemplo, componentes de las membranas.
- Alteraciones de proteínas implicadas en procesos de **transmisión y transducción de señales** (hormonas, factores de crecimiento, neurotransmisores y neuromoduladores, receptores de membrana o intracelulares, etc.), en la **homeostasis** o en la **defensa** (inmunidad).

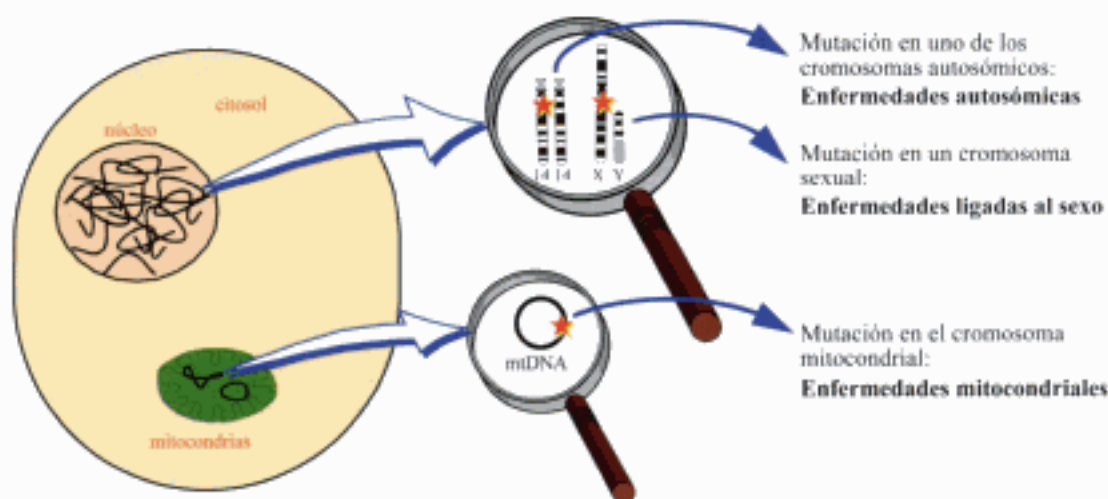
28.3.2.3 Extensión de la alteración

Aunque el tamaño de la alteración no conduce a una clasificación estricta, este criterio, ya utilizado para describir las mutaciones (pág. 344), es una simplificación adecuada a efectos pedagógicos. Se distinguen tres categorías de alteraciones genéticas, y dentro de cada una se pueden considerar otros factores, como origen de la alteración, consecuencias funcionales, mecanismo patogénico, producto génico alterado y región génica afectada.

- La **alteración puntual**, debida a la mutación de un solo nucleótido, da lugar a las **enfermedades monogénicas**, que se describen a continuación.
- Una **alteración mediana** tiene lugar normalmente en secuencias repetidas, pero no se discutirán por ser poco conocidas.
- Las **alteraciones a gran escala, cromosómicas o citogenéticas** afectan a cromosomas completos o a grandes fragmentos cromosómicos. Su incidencia se ha estimado en un 0,65%. Se distingue entre anomalías numéricas, es decir, en el número de cromosomas existentes en la célula, debidas a segregación cromosómica errónea (mutación genómica), y anomalías estructurales o cromosomopatías, debidas a reordenamientos cromosómicos. Se describirán en el tema 29.

28.3.2.4 Tipo de cromosoma afectado

Las enfermedades moleculares pueden afectar al **genoma nuclear**, caso más frecuente por su mayor tamaño, o al **genoma mitocondrial**. Dentro de las nucleares, el defecto genético puede estar localizado en los **cromosomas sexuales** (1 par) o en los **cromosomas autosómicos** (22 pares). Esta distinción, combinada con el carácter dominante o recesivo con el que se expresa la enfermedad, es la clasificación más sencilla, la más comúnmente empleada en genética, y se utilizará para describir la transmisión por herencia de las enfermedades monogénicas (pág. 394).



28.3.3 Enfermedades monogénicas nucleares

Se pueden considerar como modelo de enfermedades genéticas, al ser las que se han estudiado mejor y con mayor rigor a escala molecular. Se deben a una mutación puntual que afecta a un único gen, originando un alelo anormal. La discusión se centrará en las alteraciones localizadas en el genoma nuclear, dejando el caso de las mitocondriales para un apartado específico. Si el alelo anormal aparece en los dos cromosomas homólogos, el individuo es homocigótico para dicho gen; si aparece sólo en uno, con el alelo normal en el otro, el individuo es heterocigótico.

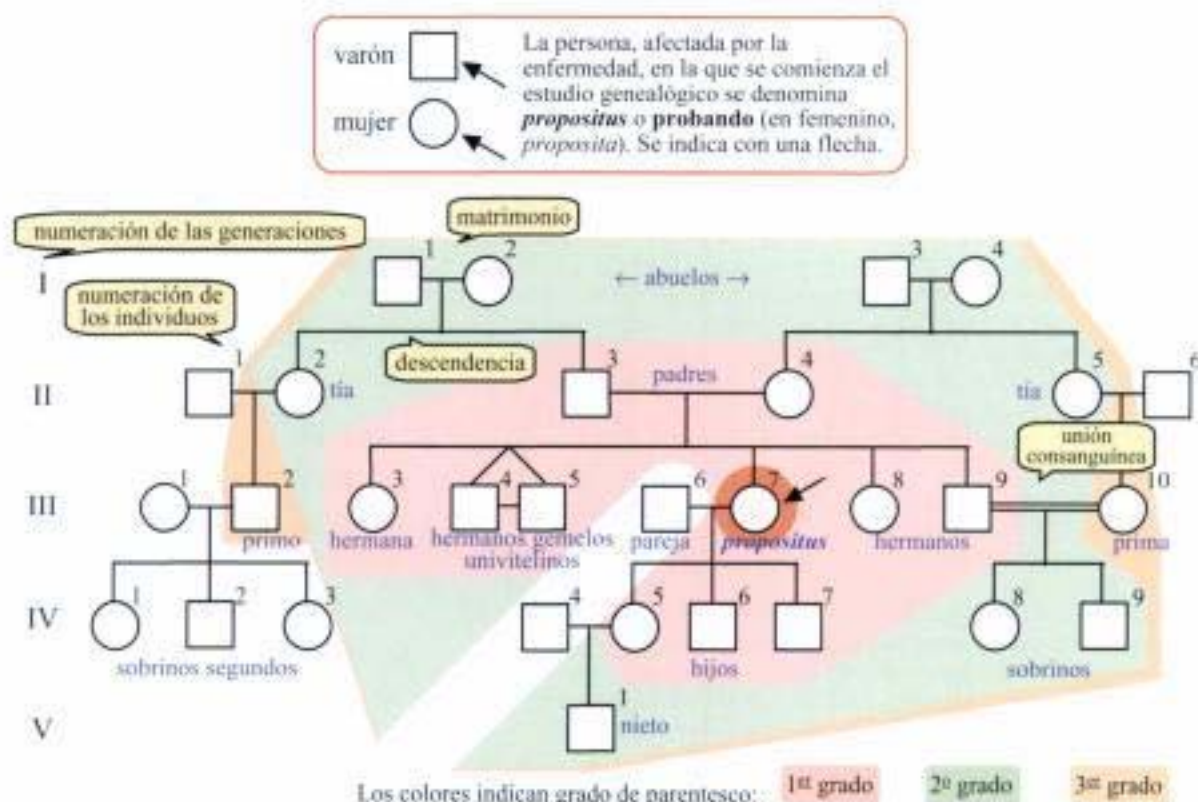
28.3.3.1 Ejemplos de enfermedades monogénicas

Las alteraciones monogénicas generan muchas enfermedades. A título meramente informativo pueden citarse acondroplasia; albinismo; anemia falciforme (o drepanocitemia); daltonismo (ceguera para algunos colores); deficiencias enzimáticas de adenosina-desaminasa, α_1 -antitripsina (pág. 376), glucosa-6-fosfato-deshidrogenasa, galactosa-1-fosfato uridiltransferasa (o UTP-hexosa-1-fosfato-uridiltransferasa, causante de la galactosemia, pág. 373) y fenilalanina-4-monooxigenasa (o fenilhidroxilasa, responsable de la fenilcetonuria); distrofia miotónica; distrofia muscular de Duchenne; corea de Huntington; enfermedades de Charcot-Marie-Tooth, de Tay-Sachs, de von Gierke y de Wilson; mucoviscidosis (o fibrosis quística, pág. 351); hemofilias; hipercolesterolemia familiar; neurofibromatosis de tipo I (págs. 352 y 395); osteogénesis imperfecta; retinoblastoma (pág. 430); síndromes de Hunter y Hurler (mucopolisacaridosis II y I); síndrome del X frágil; tumor de Wilm, etc. La información actualizada y más completa sobre caracteres mendelianos humanos (fisiológicos y patológicos) se encuentra en el catálogo de McKusick (Mendelian Inheritance in Man), publicado cada 2 años, y se centraliza en la base de datos en internet OMIM (Online Mendelian Inheritance in Man); en ambos se le asigna a cada gen un número de 6 dígitos.

28.3.3.2 Transmisión en familias

Puesto que los trastornos monogénicos dependen del genotipo en un solo locus, sus patrones de herencia familiar son muy sencillos; se llaman **patrones mendelianos** porque se ajustan a lo descrito por la 1ª ley de Mendel (ley de la segregación). Los dos homólogos de cada par de cromosomas del gametocito primario ($2n$, $4C$) se separan o segregan en la meiosis I (pág. 104), de forma que cada cromosoma homólogo aparece en un gameto secundario (n , $2C$). En la meiosis II se separan al azar las dos cromátidas para formar los gametos (n , C). Globalmente ha ocurrido una distribución independiente, aleatoria (pág. 105), de los cromosomas paternos y maternos y de sus cromátidas, con lo que cualquier individuo hereda de cada progenitor (y transmite luego a cada descendiente) un cromosoma de cada par autosómico y un cromosoma sexual. Por tanto, sólo se transmite uno de los dos alelos del gen presente en cada locus.

Para establecer los patrones de herencia de estos trastornos se acude a la información facilitada por los antecedentes familiares del paciente, que se resumen en forma de **árbol genealógico**. En éste se emplean una nomenclatura y símbolos particulares:



28.3.3.3 Clasificación de los trastornos monogénicos

Como se ha comentado (pág. 392), es común clasificar las enfermedades monogénicas de acuerdo con el tipo de cromosoma afectado (sexual o autosómico) y el carácter dominante o recesivo (pág. 95) de su expresión. Si el locus patológico está situado en un cromosoma sexual, se habla de **enfermedad ligada al sexo**; en caso contrario, de **enfermedad autosómica**.

Los caracteres autosómicos y los ligados al cromosoma X se pueden manifestar de forma dominante o recesiva, mientras que los ligados al cromosoma Y no tienen esa doble posibilidad, pues sólo existe un alelo en cada individuo, que ha de manifestarse por sí mismo, sin interacción con un homólogo.

Recordatorio de dominancia de alelos (véase pág. 95):

- Un alelo es **dominante** sobre otros cuando su presencia se manifiesta siempre en el fenotipo, independientemente de cuál sea el otro alelo (uno paterno y el otro materno). El fenotipo es igual para los individuos homocigóticos del dominante (alelos idénticos en ambos homólogos) que para los heterocigóticos (alelos diferentes en cada uno). En éstos se anula, pues, la funcionalidad del otro alelo, el **recesivo**.
- Dos alelos son **codominantes** (o **semi-dominantes**) cuando el fenotipo heterocigótico refleja la presencia de ambos, siendo intermedio entre los dos fenotipos homocigóticos. Dicho de otro modo, se detecta la expresión de cada alelo en presencia del otro.

Aplicación a la clasificación de las enfermedades genéticas monogénicas:

- Se dice que **una enfermedad es dominante** o sigue un patrón de **herencia dominante** cuando el alelo patológico es dominante sobre el normal. La enfermedad (fenotipo dominante) se manifiesta igual en homocigotos que en heterocigotos; basta un alelo patológico para que se presente la enfermedad. Como ejemplo de enfermedad autosómica dominante, igualmente grave en el genotipo homocigótico y el heterocigótico, puede citarse la corea de Huntington.
- Las enfermedades de carácter semi-dominante o codominante se manifiestan de forma más leve en los heterocigóticos que en los homocigóticos.
- Las **enfermedades recesivas** o con **herencia recesiva** están causadas por un alelo recesivo frente al normal. Sólo se manifiesta la enfermedad cuando los dos alelos son anormales, es decir, en homocigotos.

Frecuencia de las enfermedades monogénicas

- La incidencia es de 1 caso por cada 100 nacimientos vivos
- El 70% son dominantes, el 2,5% recesivas, 0,4% ligadas al cromosoma X y muy escasas las ligadas al Y

A continuación se presentan los 5 tipos posibles de enfermedad monogénica, de acuerdo con su patrón de transmisión hereditaria. Se ilustran mediante árboles genealógicos, en los cuales se han utilizado los siguientes criterios unificadores, bajo un planteamiento pedagógico:

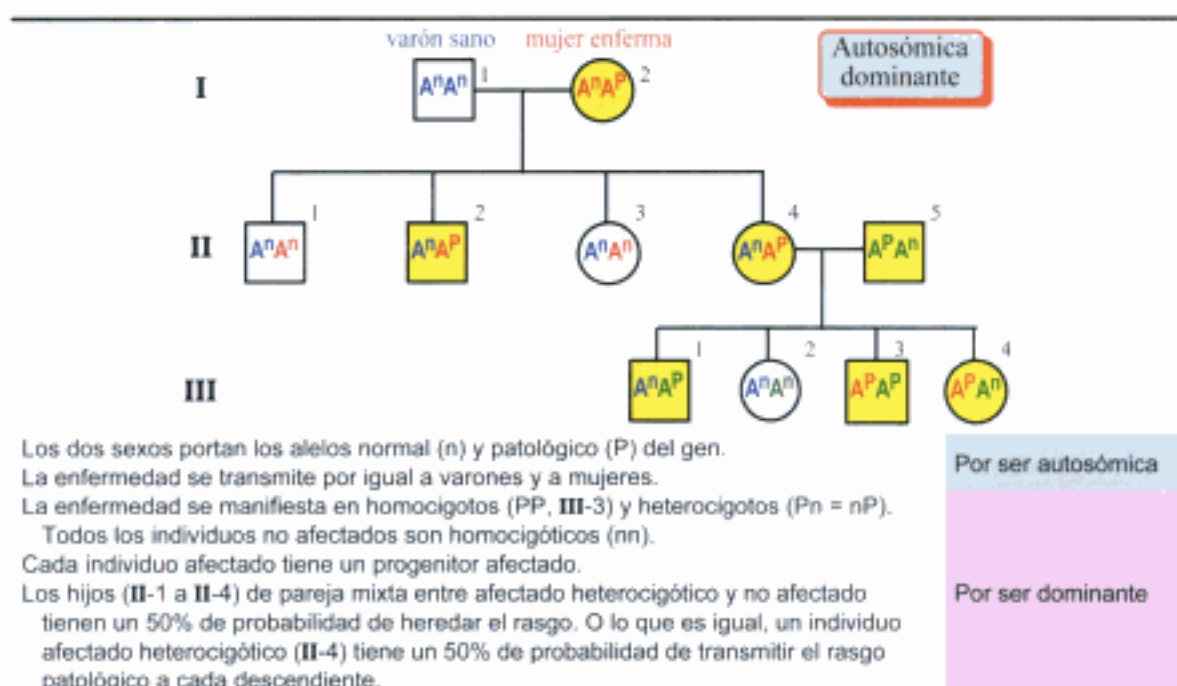
- Se consideran 3 generaciones (**I**, **II** y **III**) y sólo dos parejas (en las generaciones **I** y **II**; el 5º árbol es una excepción obligada). Ello supone, obviamente, una limitación en cuanto a posibilidades de emparejamiento y consiguiente aparición de individuos sanos o enfermos, homocigóticos o heterocigóticos, varones o mujeres. Entendemos que la exhaustividad de presentación de posibilidades de transmisión, al igual que el uso de datos familiares reales, va en detrimento de la claridad conceptual que se persigue.
- La generación **I** está formada por una pareja cuyos dos alelos (siempre referidos a un único locus génico) se muestran en color azul para la primera persona y en rojo para la segunda. La asignación de alelos a cada persona varía en los 5 árboles, para adecuarse a lo que se quiere ilustrar en cada caso.
- La generación **II** está formada por 4 descendientes, aquéllos resultantes de la combinación de los 4 alelos progenitores. De los 4 hijos, 2 son varones y 2 mujeres.
- La generación **III** está formada por los 4 descendientes resultantes de la combinación de alelos de la pareja formada por uno de los individuos de la generación **II** y una nueva persona.
- Para facilitar el seguimiento de la herencia, el genotipo de cada individuo se expresa con los siguientes colores y letras:

- Alelos del primer miembro de la pareja **I**.
- Alelos del segundo miembro de la pareja **I**.
- Alelos del nuevo miembro de la pareja **II** (el ajeno a la familia inicial).
- Letra **N** (o **n**) para el alelo normal y **P** (o **p**) para el patológico.
- Letra **A** para alelos presentes en autosomas.
- Letras **X** e **Y** para alelos en cromosomas sexuales.
- Letra mayúscula (superíndice) para alelos dominantes.
- Letra minúscula (superíndice) para alelos recesivos.

	varón	mujer
sano		
afectado		
heterocigótico o portador		

a) *Enfermedades autosómicas dominantes*

El gen responsable, localizado en uno de los autosomas, posee un alelo patológico (P) que es dominante sobre el normal (n).



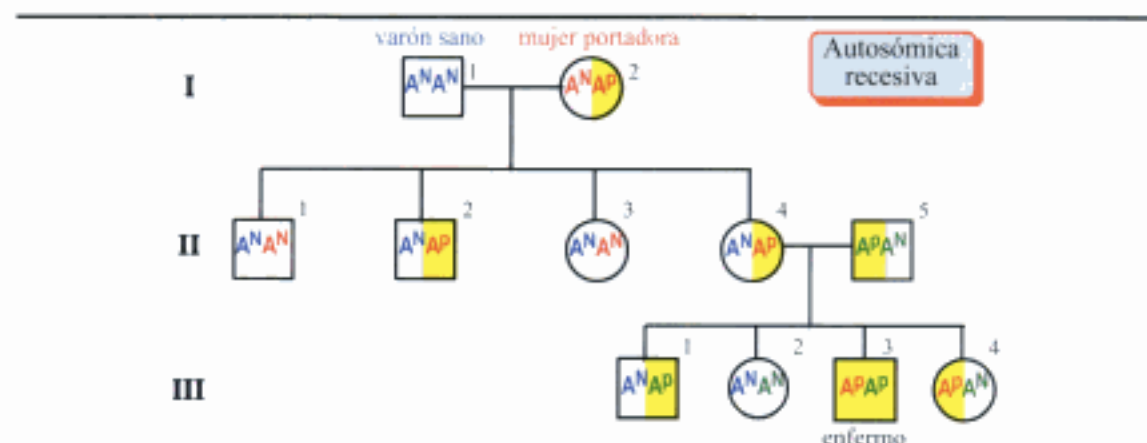
Estos trastornos constituyen cerca del 45% de las enfermedades monogénicas. Aunque muchos son poco comunes individualmente, son tan numerosos que su incidencia en conjunto es apreciable, y en ocasiones especialmente alta en áreas geográficas específicas.

Ejemplos de enfermedades autosómicas dominantes

enfermedad	proteína afectada	locus del gen	descripción
Hipercolesterolemia familiar de tipo IIA	Receptor de lipoproteínas LDL	19p13.2	
Corea de Huntington	"Huntingtina"	4p16.3	
Neurofibromatosis de tipo I	Neurofibromina	17q11.2	pág. 352
Síndrome de Marfan	Fibrilina-1	15q21.1	

b) *Enfermedades autosómicas recesivas*

En este caso, el gen responsable, localizado en uno de los autosomas, posee un alelo patológico (p) que es recesivo con respecto al normal (N).



Los dos sexos portan los alelos normal (N) y patológico (p) del gen.
 La enfermedad se transmite por igual a varones y a mujeres.
 La enfermedad se manifiesta sólo en homocigotos (pp, **III-3**). Hay individuos no afectados homocigóticos (NN) y heterocigóticos (Np = pN). Estos últimos son portadores, se dice que tienen el "rasgo", no la enfermedad..
 Los padres de personas afectadas son generalmente heterocigóticos, portadores asintomáticos, clínicamente normales.
 Cada individuo afectado tiene dos progenitores afectados o portadores. Cada portador tiene al menos un progenitor portador o afectado.

Por ser autosómica

Por ser dominante

La **anemia falciforme** es un ejemplo típico de trastorno autosómico recesivo, que fue además la primera prueba experimental de una enfermedad causada por mutación (en los años 50). El síndrome clínico es consecuencia de la alteración de un nucleótido en el gen de la β -globina (pág. 350), conduciendo a una proteína β^S . Los homocigotos para el gen β^S manifiestan el trastorno (**anemia falciforme**), mientras que los heterocigotos $\beta^A\beta^S$ se dice que muestran el **rasgo falciforme**: producen tanto hemoglobina normal (A) como anómala (S), una parte de sus eritrocitos son falciformes y padecen una anemia leve. Por tanto, la dominancia es incompleta. El diagnóstico es sencillo mediante RFLP (pág. 382).

Otras enfermedades autosómicas recesivas

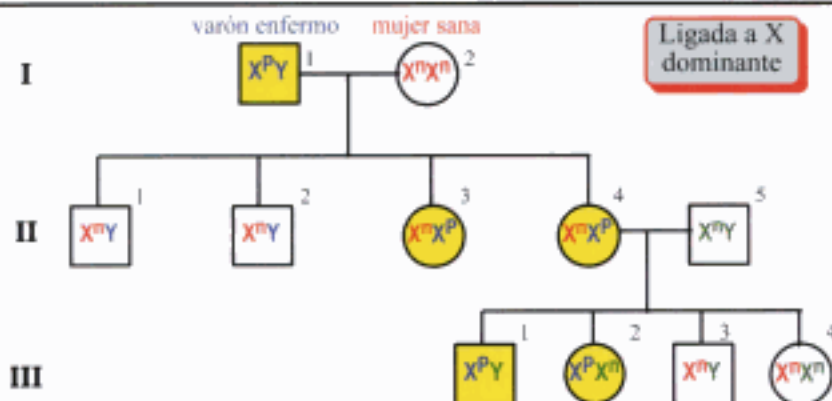
enfermedad	proteína afectada	locus del gen	descripción
Deficiencia de ADA	Adenosina-desaminasa	20q13.11	
Enfisema hereditario	α_1 -Antitripsina	14q32.1	pág. 376
Mucoviscidosis (fibrosis quística)	Regulador de la conductancia transmembrana (CFTR)	7q31.2	pág. 351
Enfermedad de Tay-Sachs	Subunidad α de la isoenzima A de la β -N-acetilhexosaminidasa	15q23-q24	pág. 350
Fenilcetonuria	Fenilalanina-4-monooxigenasa	12q24.1	
α -Talasemia	α -Globina	16pter-p13.3	
β -Talasemia	β -Globina	11p15.5	

En conjunto, este tipo de trastornos representan alrededor del 25% de las enfermedades monogénicas.

c) Enfermedades dominantes ligadas al cromosoma X

Las mujeres pueden ser homocigóticas o heterocigóticas para un gen que esté situado en un locus del cromosoma X, pues tienen dos copias de este cromosoma sexual. No ocurre así en los varones, que, al poseer un solo X, no pueden ser homocigóticos ni heterocigóticos para ese gen o rasgo; se dice que el varón es **hemigigótico**.

Si el alelo patológico (P) es dominante sobre el normal (n), las mujeres padecerán la enfermedad tanto cuando posean uno como los dos alelos P, mientras que los varones la padecerán si su único alelo es P.



Los dos sexos portan los alelos normal (n) y patológico (P) del gen, pero las mujeres portan 2 copias y los varones sólo una.

La enfermedad se transmite a varones y mujeres, pero afecta con mayor frecuencia a éstas.

Las mujeres afectadas (**II-4**) transmiten la enfermedad por igual a hijos e hijas (**III-1 a III-4**), con una probabilidad del 50% si son heterocigóticas (Pn) y del 100% si fueran homocigóticas (PP). Por el contrario, los varones afectados (**I-1**) sólo la transmiten a sus hijas (**II-3 y II-4**).

La enfermedad se manifiesta en mujeres homocigóticas (PP) y heterocigóticas (Pn), y en varones hemigigóticos (P-). Los individuos no afectados son mujeres homocigóticas (nn) y varones hemigigóticos (n-).

Cada individuo afectado tiene un progenitor afectado.

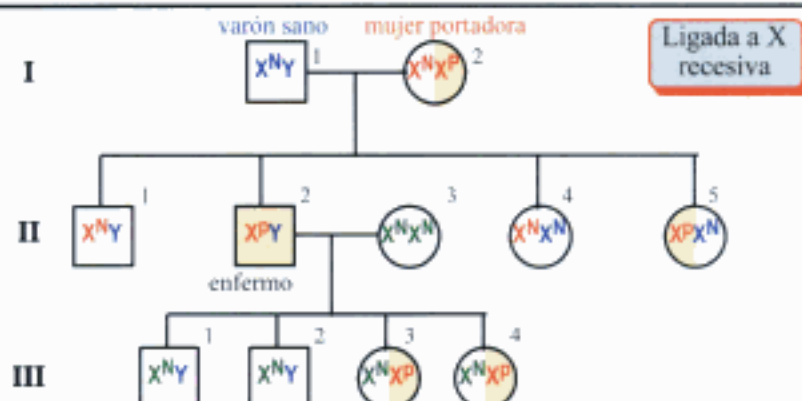
Por ser ligada al X

Por ser dominante

Para compensar la presencia de dos copias del gen en las mujeres ("doble complemento"), sus células sólo expresan los alelos de un cromosoma X (el otro está inactivado como cromatina facultativa, pág. 87). En consecuencia, las mujeres heterocigóticas pueden estar o no afectadas, dependiendo de que se exprese el alelo patológico o el normal; generalmente, esta manifestación es incluso diferente en unas células y otras, pues la selección de cuál de los cromosomas X resulta inactivado tiene lugar durante una etapa embrionaria temprana pero multicelular y es aleatoria en cada una de sus células; una vez establecida la inactivación, todas las células descendientes posteriores mantienen la selección.

d) Enfermedades recesivas ligadas al cromosoma X

Al ser en este caso dominante el alelo normal (N) sobre el patológico (p), las mujeres padecen la enfermedad sólo cuando son homocigóticas, y por tanto con frecuencia muy baja. Los varones la padecerán de forma similar al caso anterior, cuando posean el alelo patológico.



Los dos sexos portan los alelos normal (N) y patológico (p) del gen, pero las mujeres portan dos copias y los varones sólo una.
 La enfermedad se trasmite a varones y mujeres, pero afecta con mayor frecuencia a aquéllos.
 Las mujeres portadoras (I-2) transmiten la enfermedad por igual a hijos e hijas (II-1, 2, 4 y 5), con una probabilidad del 50%. Las mujeres enfermas (pp) también, pero con un 100%.
 Por el contrario, los varones enfermos (II-2) sólo la transmiten a sus hijas (III-3 y III-4).
 La enfermedad se manifiesta en mujeres homocigóticas (pp) y varones hemicigóticos (p-).
 Los individuos no afectados son mujeres homocigóticas (NN) o heterocigóticas (Np) y varones hemicigóticos (N-).
 Los varones afectados tienen progenitores sanos; la madre es generalmente portadora asintomática.

Por ser ligada al X

Por ser dominante

Como ejemplo característico, la **hemofilia A** se debe a un defecto en el factor VIII de la coagulación. El gen de esta proteína está situado en el cromosoma X (banda Xq28), por lo que la transmisión sigue el patrón ligado al sexo. Se han identificado varias alteraciones distintas del gen que causan la enfermedad. Sus alelos se expresan de forma recesiva con respecto al alelo normal del factor VIII, por lo que las mujeres sufren hemofilia con mucha menor frecuencia que los varones (se suele decir que la hemofilia A sólo la padecen los varones y no las mujeres, pero esto no es exacto, como se discute a continuación). La frecuencia de alelos patológicos en la población es de alrededor de uno por cada 10.000 alelos normales, por lo que la incidencia en varones (hemicigóticos para el gen) es de 1/10.000. Las mujeres sólo padecen la enfermedad cuando son homocigóticas para alelos patológicos, por lo que la incidencia sería de 1/10⁸; con una población mundial de 6.000 millones, esto equivale a 30 casos vivos de mujeres afectadas en todo el mundo, frente a 300.000 varones; de ahí que se asuma que las mujeres no padecen la enfermedad. En cualquier caso, si es relevante la frecuencia de mujeres portadoras, heterocigóticas para el gen del factor VIII, que pueden transmitir la enfermedad a sus hijos varones (1/10.000). Sólo la descendencia de una mujer portadora ($X^N X^p$) con un varón hemofílico ($X^p Y$) puede dar lugar a mujeres enfermas ($X^p X^p$).

Otros ejemplos de enfermedades recesivas ligadas al cromosoma X son la distrofia muscular de Duchenne (Xp21.2) y el daltonismo (Xq28).

e) Enfermedades ligadas al cromosoma Y

Se expone este caso por razones pedagógicas, a pesar de que no se conoce ninguna enfermedad ligada a Y. La primera razón es el pequeño tamaño del cromosoma Y (en realidad, aún menor, pues parte del Y tiene región homóloga en el X y esos loci se comportan como los autosómicos); por ello, se conocen escasos ejemplos de caracteres asociados a Y, y ninguno es patológico. La segunda razón es que para que exista una patología debe haber un gen cuya función sea importante y, teniendo en cuenta que las mujeres carecen de tales locus ligados al Y, no podrían desarrollar esa función.

Entre las enfermedades asociadas a mutaciones del DNA mitocondrial se pueden citar las siguientes:

- Neuropatía óptica hereditaria de Leber (*NOHL*). Fue la primera enfermedad donde se demostró la herencia por vía materna de un defecto en el genoma mitocondrial.
- Oftalmoplejía externa progresiva crónica (*CPEO*).
- Síndrome de Kearns-Sayre (*KSS*), una oftalmoplejía externa con complicaciones varias.
- Epilepsia mioclónica con fibras rojas de aspecto raído (*MERRF*).
- Encefalomiopatía mitocondrial con acidosis láctica y episodios similares a apoplejía (*MELAS*).

Adicionalmente, es probable la implicación de mutaciones mitocondriales somáticas o adquiridas como uno de los factores que contribuyen al proceso de envejecimiento y a enfermedades neurodegenerativas, incluyendo la enfermedad de Alzheimer.

28.4 ENFERMEDADES POLIGÉNICAS, MULTIFACTORIALES, MIXTAS O COMPLEJAS

Muchas enfermedades genéticas no están causadas únicamente por el defecto de un gen, dando lugar a un grupo de trastornos conocidos bajo nombres diferentes, de difícil delimitación o caracterizados de forma incompleta. A diferencia de las enfermedades monogénicas, la manifestación clínica es un reflejo del efecto combinado o interacción acumulativa de alteraciones producidas por factores genéticos (múltiples mutaciones, a veces simultáneas en un número impreciso de genes) y ambientales o exógenos (de todo tipo: nutricional, agentes tóxicos, estrés, etc.). En cualquier caso, la complejidad del origen explica el conjunto de efectos diversos producidos y los distintos nombres recibidos. Estas enfermedades multifactoriales no deben confundirse con los trastornos cromosómicos (enfermedades genéticas a gran escala, tema 29), aunque en ambas pueda tener lugar una alteración compleja del genoma.

Al no ser trastornos de un gen único, la herencia de estas alteraciones no sigue la genética clásica mendeliana. A pesar de ello, si se observa una mayor transmisión de estas enfermedades dentro de una familia que entre individuos no emparentados. Para que el síndrome se transmita a otro miembro de la misma familia es necesario que se reúna una combinación de cambios génicos.

Entre estas enfermedades pueden citarse varias malformaciones congénitas, como el labio leporino, las cardiopatías congénitas y los defectos del tubo neural (anencefalia, espina bífida y otras), y ciertas enfermedades de la edad adulta, como algunos tipos de cáncer y las enfermedades de las arterias coronarias, donde se ha observado una predisposición genética. De todas ellas sólo se considerará el cáncer, la más estudiada desde el punto de vista molecular además de por su trascendencia clínica (tema 30).

Tema 29

Enfermedades cromosómicas o citogenéticas

29.1	Introducción	401
29.2	Clasificación y nomenclatura	402
29.3	Cromosomopatías numéricas.....	403
29.3.1	Euploidia	403
29.3.1.1	Triploidia	403
29.3.1.2	Tetraploidia.....	404
29.3.2	Aneuploidia	404
29.3.2.1	Mecanismo de generación de la aneuploidia	404
29.3.2.2	Monosomía	404
29.3.2.3	Trisomía.....	406
29.3.3	Mixoploidia	406
29.4	Cromosomopatías estructurales	406
29.4.1	Un punto de ruptura en un cromosoma: deleción terminal	406
29.4.2	Dos puntos de ruptura en un cromosoma	407
29.4.2.1	Inversión	407
a)	Inversión paracéntrica.....	407
b)	Inversión pericéntrica.....	407
29.4.2.2	Deleción intersticial.....	408
29.4.2.3	Cromosomas anulares.....	408
29.4.3	Dos puntos de ruptura en cromosomas diferentes: translocación	409
29.4.3.1	Translocación recíproca	409
29.4.3.2	Translocación de Robertson o fusión céntrica	409
29.4.4	Isocromosomas	410

29.1 INTRODUCCIÓN

En este tema se tratarán las enfermedades cromosómicas, uno de los tipos de enfermedad molecular (pág. 392). Como su nombre indica, se originan por la alteración del número o la estructura de los cromosomas; de ahí que se llamen también alteraciones cromosómicas a gran escala, aberraciones, trastornos o anomalías cromosómicas. Al poder observarse los cambios con el microscopio óptico, son también consideradas como enfermedades citogenéticas.

Son alteraciones esporádicas y de baja frecuencia, que raramente se perpetúan al ser casi incompatibles con la supervivencia y la reproducción. Sin embargo, cuando lo hacen, muestran manifestaciones clínicas, físicas o psíquicas, muy graves, a veces dramáticas, por lo que pueden considerarse las principales enfermedades genéticas. De hecho, son responsables de una gran proporción de abortos espontáneos, malformaciones congénitas, retrasos mentales y tumores.

Como parte de las consecuencias de estas enfermedades, aparecen también alteraciones en el metabolismo y en la síntesis o actividad de proteínas. Tales variaciones bioquímicas pueden considerarse secundarias a la enfermedad, a diferencia de las que aparecen en las enfermedades monogénicas, pero son igualmente objeto de estudio de la patología molecular actual.

Algunas cifras significativas, relativas a las anomalías cromosómicas

- Se presentan en el 4% de los nacidos muertos
- Aparecen en el 2% de los embarazos de mujeres mayores de 35 años
- Presentes en el 50% de los abortos espontáneos durante el primer trimestre
- La frecuencia es aún mayor e imprecisa en la etapa embrionaria (antes de las 8 semanas de gestación), del orden de 75 por cada 1.000 embriones
- Se han identificado unos 60 síndromes

Las alteraciones son mucho más extensas que las producidas en los trastornos monogénicos, caracterizándose por grandes cambios en el número de nucleótidos (miles o millones), en la estructura de los cromosomas o en la reasociación de genes por inserción o delección, duplicación, translocación, inversión, etc.

Pueden afectar tanto a los autosomas como a los cromosomas sexuales; en ambos casos, pueden ser constitutivas o adquiridas.

- Son **aberraciones constitutivas** aquellas que afectan a todas las células del organismo, por aparecer durante la división meiótica de las células germinales (gametogénesis), o bien durante la primera división mitótica del cigoto. Como consecuencia, son las de mayor gravedad.
- Las **aberraciones adquiridas** o somáticas sólo ocurren durante las mitosis poscigóticas. Se inician, por tanto, en células somáticas, sin afectar a todas las células del organismo; de ahí la menor gravedad (a veces mínima) de sus consecuencias.

29.2 CLASIFICACIÓN Y NOMENCLATURA

Anomalías
cromosómicas

Núméricas: se altera el número de cromosomas

Estructurales: se modifica la estructura de algún cromosoma

Entre los abortos por anomalías cromosómicas durante el primer trimestre de embarazo, el 96% se deben a alteraciones numéricas y el 4% a estructurales. Igualmente, en madres mayores de 35 años, el 85% de los abortos debidos a cromosomopatías corresponden a las anomalías numéricas y el 15% a las estructurales.

En la descripción de las anomalías cromosómicas, la **nomenclatura** utilizada se basa en la de la dotación cromosómica normal, es decir, **46,XY** en células diploides del varón y **46,XX** en la mujer. Se añaden abreviaturas para expresar aspectos relativos a la morfología del cromosoma o para indicar las causas de la anomalía. Los detalles de aplicación de la nomenclatura se describen en cada anomalía.

46,XY	célula masculina normal, diploide: 46 cromosomas, incluyendo un X y un Y	46,XX	célula femenina normal, diploide: 46 cromosomas, incluyendo dos X
47,XXY	un cromosoma X adicional	39,XXX	célula femenina triploide
p	brazo corto	pat	origen paterno
q	brazo largo	mat	origen materno
cen	centrómero	dic	cromosoma dicéntrico
ter	extremo (terminal, telómero)	der	cromosoma derivado
del	delección	i	isocromosoma
ins	inserción	r	cromosoma anular
dup	duplicación	mar	cromosoma marcador
inv	inversión	fra	sitio frágil
t	translocación	+	ganancia
rcp	translocación recíproca	-	pérdida
rob	translocación de Robertson	:	rotura
/	mosaicismo	::	rotura y unión

29.3 CROMOSOMOPATÍAS NUMÉRICAS

También llamadas simplemente **alteraciones genómicas**, implican un cambio en el número de cromosomas, bien pérdida o ganancia, pero nunca se altera la organización del cromosoma. Son las anomalías más frecuentes y se originan en gran parte durante la división meiótica (pág. 104), con lo que se transmiten tras la fecundación a todas las células del hijo o hija. En general son letales, impidiendo que se sobrepase el desarrollo embrionario, con abortos espontáneos después de la concepción.

Las anomalías cromosómicas numéricas se clasifican así:

Nombre de la anomalía cromosómica		Nº de cromosomas	Ejemplos (humanos)
Euploidia (= xn)	Nuliploidia	0	eritrocitos y plaquetas
	Haploidia (normal)	n	23,X
	Diploidia (normal)	$2n$	46,XY
	Poliploidia	Triploidia	69,XXY
		Tetraploidia	92,XXYY
Aneuploidia ($\neq xn$)	Nulisomia	$2n-2$	44,XY,-21,-21
	Monosomia	$2n-1$	45,X 45,XX-18
	Disomia (normal, no es aneuploidia)	$2n$	
	Trisomia	$2n+1$	47,XXY 45,XX,+21
	Polisomia	$2n+x$	49,XXXXX
Mixoploidia	Mezcla de dos poblaciones celulares con distinta dotación cromosómica		46,XX/47,XX,+8

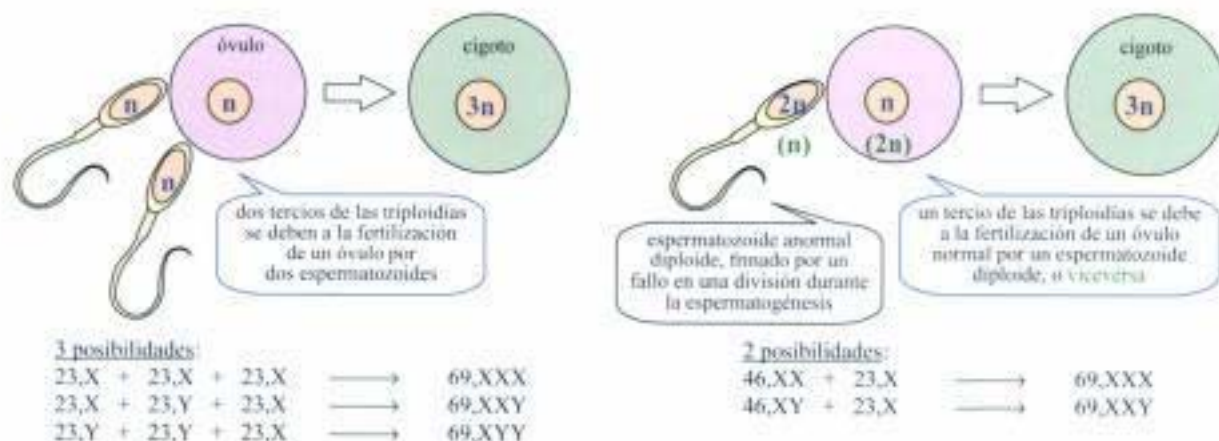
Se habla de **heteroploidia** en todas las situaciones (coloreadas) en las que el número de cromosomas es diferente al normal

29.3.1 Euploidia

Recibe este nombre (del griego *eu*, bien, verdadero: ploidia correcta o verdadera) la situación en la que existe un número entero de dotaciones cromosómicas haploides completas, es decir, el número de cromosomas es un múltiplo de n . Para los casos con más cromosomas de lo normal ($2n$) se habla de **poliploidia**. En el extremo opuesto se encuentran aquellas células diferenciadas como plaquetas y eritrocitos, que carecen de núcleo y se dice que son **nuliploides**.

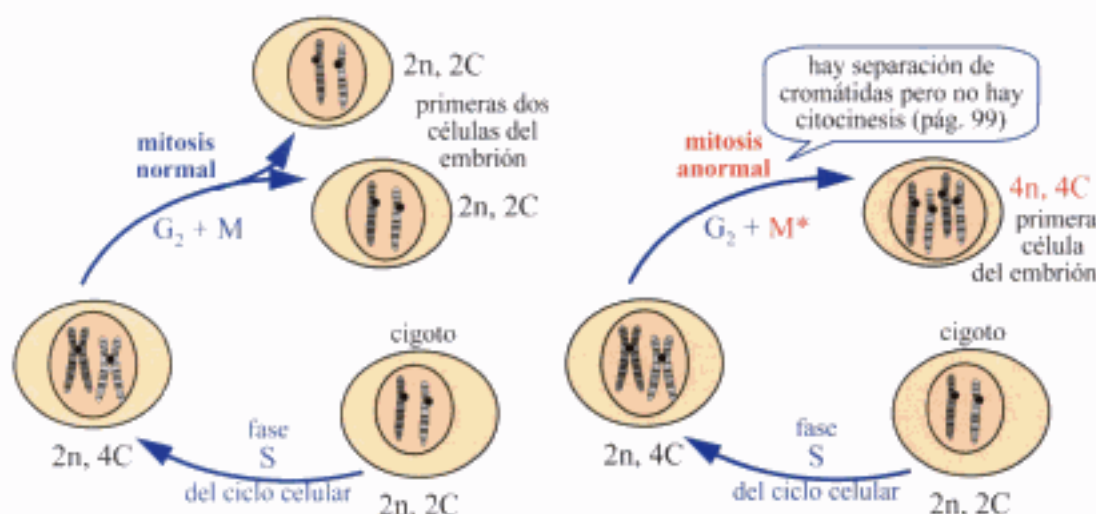
29.3.1.1 Triploidia

Es la anomalía en la que las células contienen 3 juegos cromosómicos ($3n$, en humanos, $3 \times 23 = 69$), que incluyen 3 cromosomas sexuales. Es la poliploidia más común, llegando a constituir el 2% de todas las concepciones. Sin embargo, éstas se suelen perder por aborto espontáneo, por lo que la incidencia es muy baja.



29.3.1.2 Tetraploidia

En este caso ($4n$, $4 \times 23 = 92$) la célula contiene dos dotaciones cromosómicas extras; son siempre 92,XXXX o 92,XXYY. Se origina usualmente por un fallo en la primera mitosis del cigoto: tras una replicación normal del DNA no ocurre la división celular subsiguiente, la que debería formar las dos primeras células embrionarias.



29.3.2 Aneuploidia

Este nombre se aplica a las anomalías donde se pierde el carácter euploide, es decir, el número de cromosomas no es múltiplo del número haploide n . Ello se debe a la variación del número de copias de un solo cromosoma, no de juegos completos de cromosomas. En general, se presenta una copia de más o de menos de un solo cromosoma, específico en cada alteración, produciéndose, respectivamente, una **trisomía** (3 cromosomas homólogos) o una **monosomía** (una sola copia, no hay homólogo). En consecuencia, la dotación cromosómica total es de $2n+1$ ($46+1 = 47$ cromosomas en la trisomía humana) o $2n-1$ ($46-1 = 45$ cromosomas, monosomía humana). El cromosoma extra se indica con un signo "+" y el ausente con un signo "-".

Son las anomalías numéricas más frecuentes y clínicamente significativas; aparecen en el 3-4% de los embarazos, asociándose con un desarrollo físico o mental anómalo. En las células cancerosas también aparecen a menudo aneuploidias extremas, con múltiples anomalías cromosómicas. Cuando falta un par de homólogos, la anomalía es letal, y se llama **nulismía**; por ejemplo, $44,XY,-21,-21$ (varón) o $44,XX,-21,-21$ (mujer).

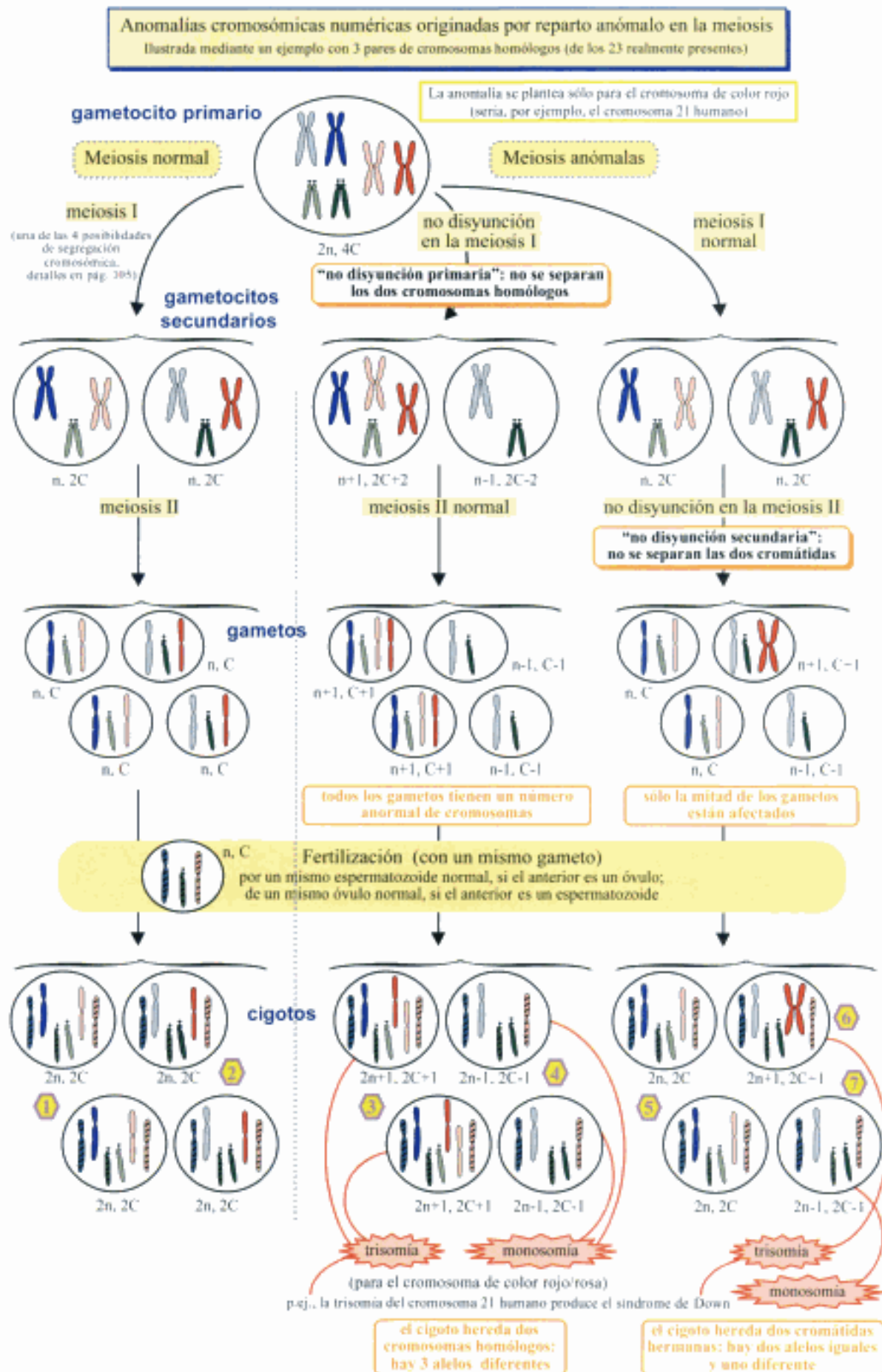
29.3.2.1 Mecanismo de generación de la aneuploidia

La aneuploidia se genera por una falta de disyunción durante la meiosis, es decir, un fallo en la separación normal de un par de cromosomas o de cromátidas. Las consecuencias del proceso son diferentes dependiendo de la división meiótica (I o II) en la que se produzca el fallo.

- La ausencia de disyunción cromosómica durante la meiosis I hace que los dos homólogos del **cromosoma afectado** pasen juntos a uno de los gametocitos secundarios, y el otro gametocito quede sin ese cromosoma. Los $2(n-1)$ cromosomas restantes han sufrido la disyunción normal y, por tanto, han separado sus homólogos y cada gametocito secundario contiene $n-1$. En la subsiguiente meiosis II, con disyunción normal, se separan todas las parejas de cromátidas. Los gametos maduros tienen una cromátida (ahora ya cromosoma hijo) de más, o bien una de menos, siempre con los otros $n-1$ cromosomas correctos.
- La no disyunción de las cromátidas del **cromosoma afectado** durante la meiosis II (tras una meiosis I normal) hace que ambas pasen a un mismo gameto, que queda así con dos copias del **cromosoma afectado**, mientras el otro no contiene ninguna; se forman también otros dos gametos normales. El resto de cromosomas ($n-1$) separan sus cromátidas correctamente.

29.3.2.2 Monosomía

Como se ha indicado, es la situación en la que hay pérdida de un solo homólogo de un par de cromosomas (en total, $2n-1$, 45 en humanos). El ejemplo más frecuente es el **síndrome de Turner**, una monosomía en cromosomas sexuales, con cariotipo $45,X$ y, por tanto, fenotipo femenino. Se origina por la no disyunción de los cromosomas sexuales (casos nº 4 o nº 7 en la figura; el cromosoma de color rosa rayado sería el X, procedente del gameto normal). Su incidencia es de 1 por cada 5.000 niñas nacidas vivas (la mayoría de las concepciones se pierden prenatalmente). Otros casos con el mismo fenotipo son **mosaicos** $45,X/46,XX$ o $45,X/46,XY$ (definidos a continuación, pág. 406).



29.3.2.3 Trisomía

En este caso existe un homólogo adicional de un cromosoma, es decir, 3 copias de dicho cromosoma (en total, $2n+1$, 47 en humanos). Puede afectar a cualquiera de los 23 pares. Existen varios ejemplos moderadamente comunes.

El **síndrome de Down** o **mongolismo** presenta (en el 95% de los casos) una copia extra del cromosoma 21 (**trisomía 21**), por lo que existen tres homólogos de este cromosoma. La nomenclatura es 47,XY,+21 para los varones y 47,XX,+21 para las mujeres. Se origina por la no disyunción del cromosoma 21 (casos nº 3 o nº 6 en la figura; los cromosomas de color rosa y rojo serían los homólogos 21). Es el tipo de trisomía más común, con una incidencia de uno por cada 800 nacidos vivos. El riesgo (también de las otras trisomías en general) se relaciona con la edad avanzada de la madre, probablemente por un aumento de la no disyunción en las mujeres mayores, de forma que la frecuencia de trisomía 21 es de 1/400 a los 35 años y 1/100 a los 40 años. Presenta, entre otros síntomas, retraso mental moderado; la supervivencia es de unos 40 años. No existe anomalía alguna a nivel génico.

El **síndrome de Edward** es la **trisomía 18**, representada 47,XX,+18 (varón) y 47,XY,+18 (mujer), con una incidencia de uno por cada 6.000 nacidos vivos.

El **síndrome de Patau** o **trisomía 13** tiene una incidencia de uno de cada 10.000 nacidos vivos. Pueden sobrevivir a término, pero muestran malformaciones congénitas múltiples y graves. La nomenclatura es 47,XY,+13 (varón) y 47,XX,+13 (mujer).

El **síndrome de Klinefelter**, con una incidencia de uno por cada 1.000 nacidos vivos, es una trisomía en cromosomas sexuales que produce varones 47,XXY. Con una incidencia parecida se encuentran también trisomías 47,XXX (mujer) y 47,YYY (varón).

29.3.3 Mixoploidía

Recibe este nombre la situación en la que el organismo contiene células con distinto número de cromosomas. Aquéllas constituyen dos (o más) líneas celulares, cada una derivada de una célula madre. Se puede distinguir entre **mosaico**, cuando las dos líneas derivan de un solo cigoto, generalmente por mutación en una etapa inicial del embrión, y **quimera**, cuando el origen son dos células madre independientes que se han unido en un solo embrión. Como ejemplo, 46,XX/47,XX,+8 indica una mujer con dos poblaciones celulares, una con cariotipo normal y otra con trisomía 8.

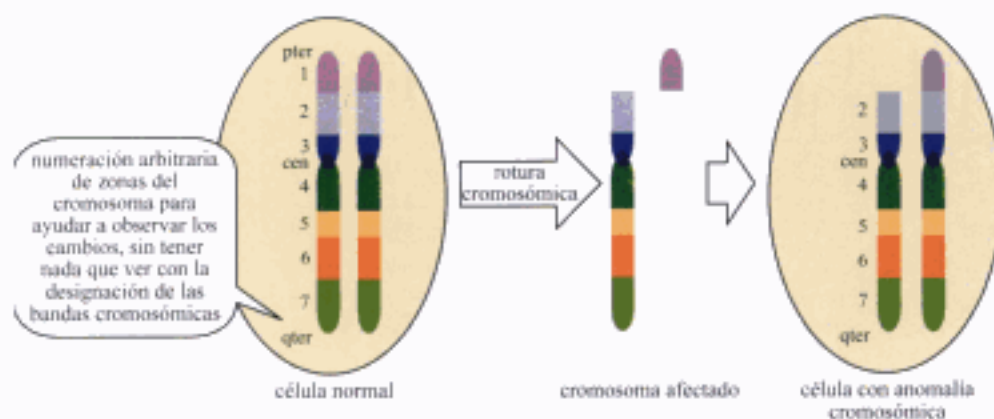
29.4 CROMOSOMOPATÍAS ESTRUCTURALES

Este segundo tipo de anomalías cromosómicas, menos frecuentes que las numéricas, corresponde a cambios en la estructura de uno o varios cromosomas, sin que se altere su número (46 en el ser humano). El cambio se debe a la rotura del cromosoma, comúnmente seguida de una reconstitución de la molécula en una disposición diferente, anormal. Son particularmente frecuentes en la anemia de Fanconi, la ataxia telangiectasia y el síndrome de Bloom.

La descripción de los diferentes tipos de anomalías estructurales tiene en cuenta el **número de puntos de rotura** del cromosoma y el carácter **equilibrado** (no hay pérdida de material cromosómico) o **desequilibrado** (se pierde parte del cromosoma).

29.4.1 Un punto de ruptura en un cromosoma: delección terminal

Normalmente, tras la ruptura de la cadena de DNA los extremos resultantes se suelen unir rápidamente por la acción de enzimas de reparación. Si la rotura ocurre en un solo punto y no se repara, el fragmento que no contiene el centrómero se perderá en la siguiente división celular. Se origina así un cromosoma con delección terminal.



Los efectos fenotípicos suelen ser graves, pues hay pérdida de material genético (anomalía **desequilibrada**). Además, aunque el cromosoma fragmentado remanente contiene un centrómero, la ausencia de telómero en el extremo resultante del corte hace que el cromosoma sea inestable y se pueda degradar fácilmente.

En cuanto a la nomenclatura, se utiliza la terminología establecida internacionalmente (pág. 90). Debe indicarse, en este caso, los números de cromosoma afectado, brazo, banda, sub-banda y sub-sub-banda donde se produjo la rotura. Por ejemplo, $46,XY,del(4)(p16.3) \rightarrow pter$ indica una delección en el brazo 4p desde la banda 16.3 hasta el extremo (*ter*) o, lo que es igual, una ruptura en la banda 1, sub-banda 6, sub-sub-banda 3 del brazo corto del cromosoma 4 de un varón.

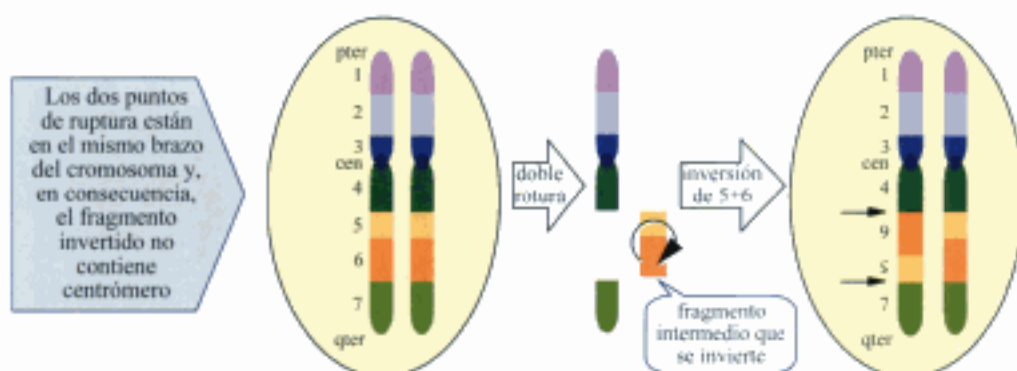
29.4.2 Dos puntos de ruptura en un cromosoma

Cuando se ha producido más de un corte, las enzimas de reparación tienen dificultades para distinguir entre los extremos resultantes y pueden unir extremos que no se correspondían en el cromosoma original. Como consecuencia, surgen también aberraciones estructurales.

29.4.2.1 Inversión

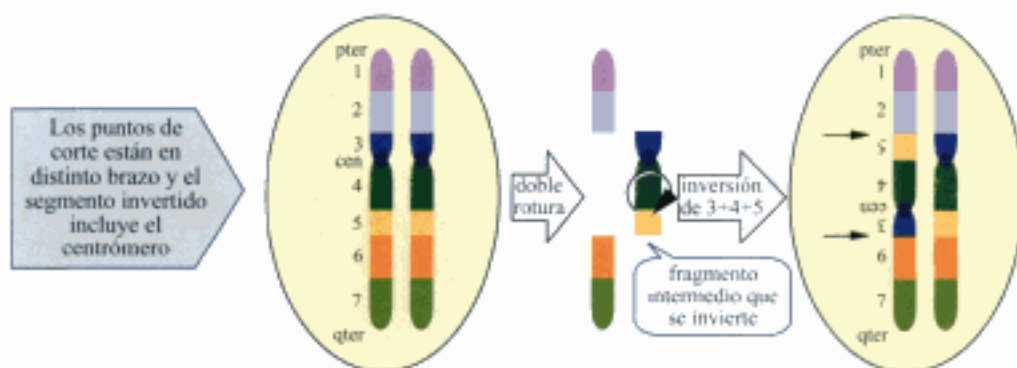
Tiene lugar cuando el fragmento intermedio formado por las dos roturas se vuelve a unir en su posición original, pero invertido. La secuencia del cromosoma resulta, por tanto, alterada. Son reordenamientos equilibrados, pues el material perdido se vuelve a incorporar y no hay pérdida neta de material cromosómico. Pueden distinguirse dos casos, con distinto resultado:

a) Inversión paracéntrica



Se nombran, por ejemplo, como $46,XY,inv(11)(p12p14)$ para una inversión entre las bandas 12 y 14 del brazo corto del cromosoma 11 de un varón; puesto que ambas posiciones son del brazo p, la inversión es paracéntrica, aunque no se indica explícitamente este carácter.

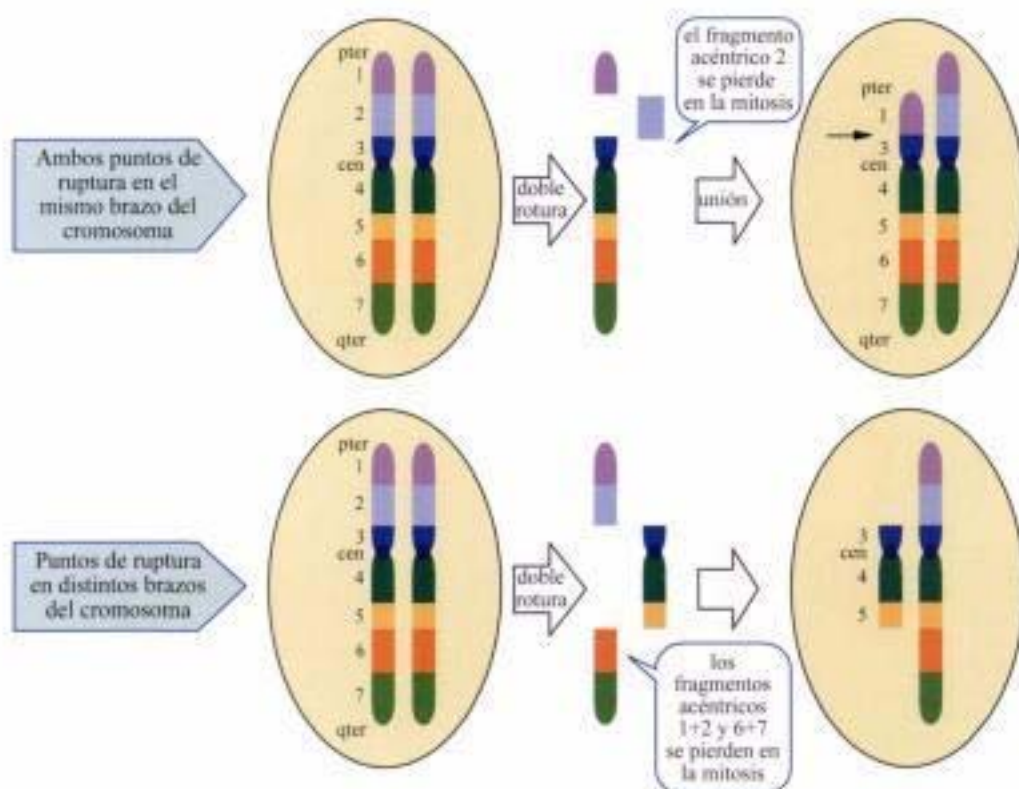
b) Inversión pericéntrica



La nomenclatura es análoga al caso anterior; por ejemplo, $46,XY,inv(3)(p21q13)$ indicando inversión en el cromosoma 3 de un varón, desde la banda 21 del brazo corto a la 13 del brazo largo (y, por tanto, pericéntrica por incluir el centrómero).

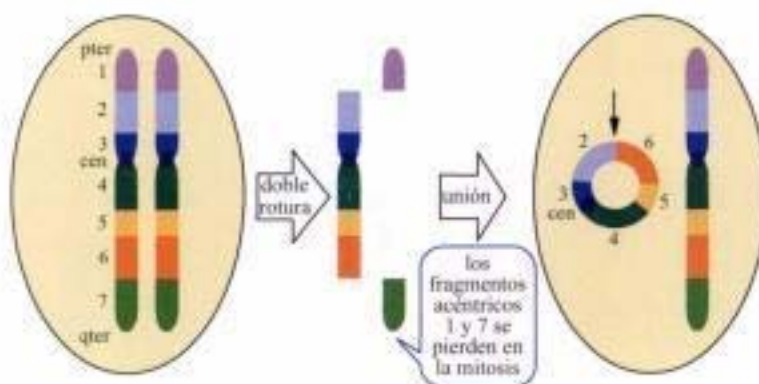
29.4.2.2 Delección intersticial

El mismo fragmento cromosómico intermedio de los ejemplos anteriores puede perderse, en lugar de volverse a insertar, por unión de los fragmentos restantes. En este caso, el reordenamiento es desequilibrado, pues hay pérdida neta de material cromosómico. Sólo es estable el fragmento (intersticial o unión de los laterales) que contiene el centrómero; el otro (fragmento **acéntrico**, pág. 92) se perderá en la siguiente división celular. Un ejemplo de nomenclatura sería 46,XY,del(7)(q31)-(q33).



29.4.2.3 Cromosomas anulares

Aparecen al unirse los dos extremos del fragmento intermedio de cromosoma eliminado en la delección intersticial. Sólo si contienen el centrómero pueden transmitirse en la división celular.

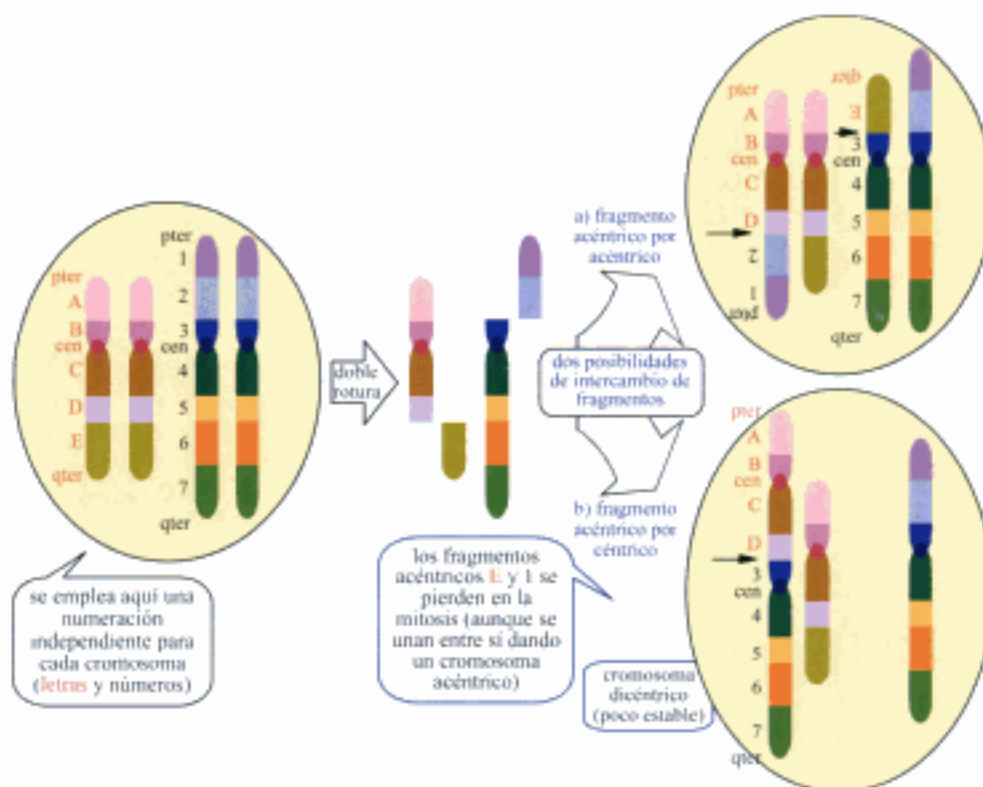


29.4.3 Dos puntos de ruptura en cromosomas diferentes: translocación

La escisión de dos cromosomas no homólogos, en un punto cada uno, puede conducir a la unión incorrecta de fragmentos entre ambos, generando cromosomas anormales "híbridos" o **derivados**.

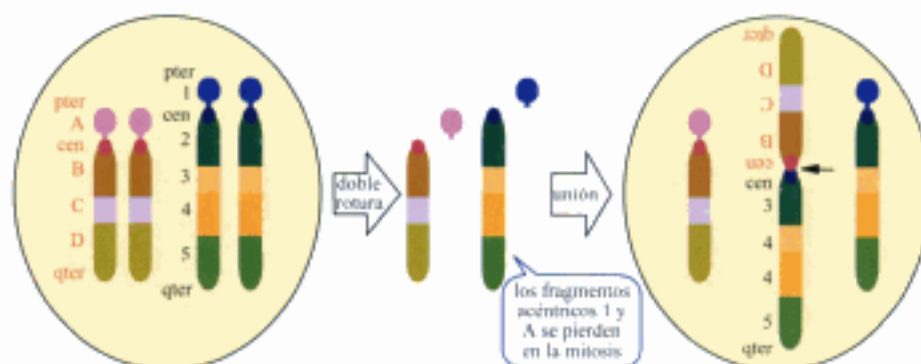
29.4.3.1 Translocación recíproca

Esta translocación implica el intercambio de material entre los dos cromosomas; al no haber pérdida de material genético en su conjunto, se trata de una anomalía equilibrada. Son anomalías relativamente frecuentes, aunque generalmente no tienen efecto fenotípico, pues los genes simplemente cambian de cromosoma.



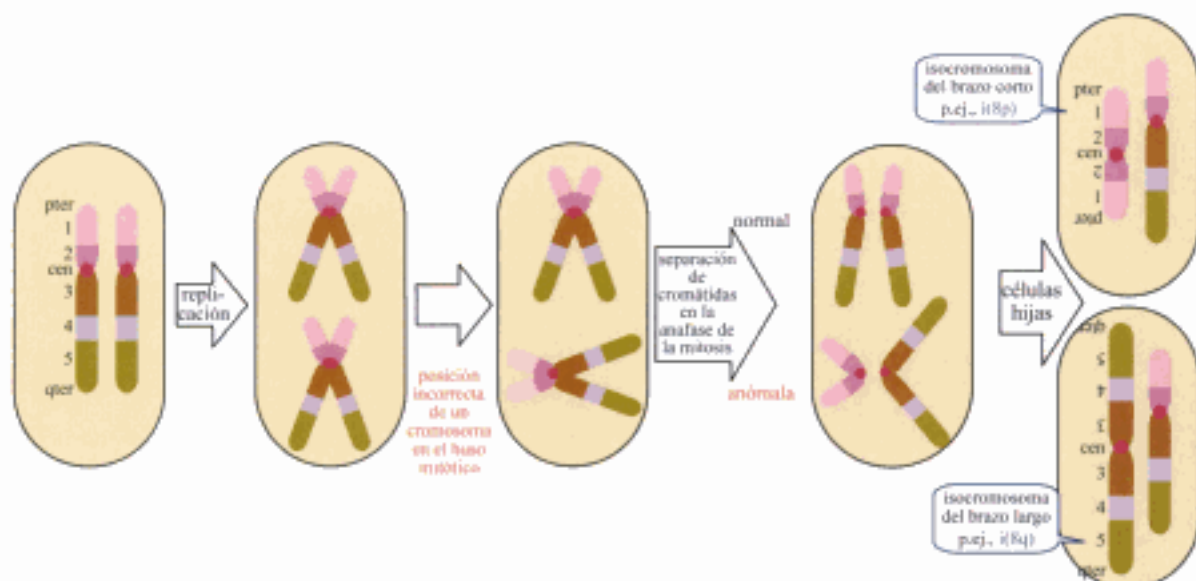
29.4.3.2 Translocación de Robertson o fusión centrómica

Es un reordenamiento particular, entre dos cromosomas acrocéntricos no homólogos, que se escinden muy cerca del centrómero y cuyos fragmentos grandes se unen por los centrómeros. Los dos fragmentos cortos, acéntricos, se pierden en la primera mitosis.



29.4.4 Isocromosomas

Como ejemplo ilustrativo de la existencia de otros tipos de anomalías cromosómicas estructurales, se mencionan los *isocromosomas*, que resultan de una división aberrante de cromosomas por el eje perpendicular al eje normal de división; las cromátidas del nuevo cromosoma están formadas por los brazos cortos o los largos del cromosoma original.



Tema 30

Bases moleculares del cáncer

30.1	Introducción	412
30.2	El cáncer como enfermedad genética	412
30.3	Etapas características en el desarrollo del cáncer	413
30.3.1	Tumor primario benigno	413
30.3.2	Cáncer <i>in situ</i>	414
30.3.3	Tumor maligno o cáncer propiamente dicho	414
30.3.3.1	Escape celular o invasividad	414
30.3.3.2	Vascularización o angiogénesis	414
30.3.3.3	Formación del tumor secundario o metástasis	414
30.4	Mecanismos de transformación de célula normal en tumoral	414
30.4.1	Equilibrio entre proliferación y muerte celular	414
30.4.2	Genes responsables del cáncer	416
30.4.2.1	Mutación de protooncogenes: oncogenes	416
a)	Concepto	416
b)	Tipos de protooncogenes	416
c)	Características de la expresión de los oncogenes	417
30.4.2.2	Mutación de genes oncosupresores	418
a)	Concepto	418
b)	Tipos de genes oncosupresores	418
c)	Características de la expresión de los genes oncosupresores	419
30.5	Control de la proliferación por señales extracelulares	420
30.5.1	Tipos de señal	420
30.5.2	Transducción de la señal extracelular al interior	421
30.6	Regulación del ciclo celular	422
30.6.1	Puntos de control en el ciclo de división celular	422
30.6.2	Proteínas que controlan el ciclo celular	424
30.6.2.1	Ciclinas	424
a)	Ciclinas de la fase G ₁ o ciclinas de inicio	424
b)	Ciclinas mitóticas o de la fase G ₂	424
30.6.2.2	Proteína quinasas dependientes de ciclinas (Cdk)	424
30.6.2.3	Actividad de los complejos Cdk-ciclina	426
30.6.3	Regulación del ciclo celular por los complejos Cdk-ciclina y alteraciones en el cáncer	427
30.6.3.1	Estímulos de progresión por el ciclo celular y oncogenes que los alteran	428
30.6.3.2	Estímulos de detención del ciclo celular	429
30.6.3.3	El gen supresor del retinoblastoma (<i>Rb</i>)	430
30.6.3.4	El gen oncosupresor <i>p53</i>	431
30.7	Apoptosis, muerte celular programada o suicidio celular	431
30.7.1	Etapas y control del proceso de apoptosis	432
30.7.1.1	Inducción	432
30.7.1.2	Ejecución	432
a)	Proteasas: caspasas y su papel en el control de la apoptosis	432
b)	Nucleasas	433
30.7.1.3	Fase final de la apoptosis	433
30.7.2	Relación entre ciclo celular y apoptosis: <i>p53</i> como punto de enlace	434
30.8	Herencia del cáncer	434
30.9	Marcadores tumorales	434
30.9.1	Alteraciones locales y generalizadas	435
30.9.2	Alteraciones tumorales y tisulares	435

30.9.2.1	Marcadores tumorales	435
a)	Marcadores tumorales de secreción o marcadores "clásicos"	435
b)	Citocinas o mediadores solubles de origen celular	437
30.9.2.2	Marcadores tisulares	437
30.9.3	Marcadores genotípicos o moleculares	437

30.1 INTRODUCCIÓN

El cáncer, neoplasia o tumor puede considerarse una enfermedad genética (véase la clasificación molecular de las enfermedades en pág. 390), que se desarrolla en organismos superiores, en la mayoría de los tejidos y en todo tipo de células somáticas (como excepción, también en las células germinales en cánceres hereditarios, pág. 434). Bajo el nombre genérico de cáncer se engloba un conjunto de enfermedades que tienen en común un crecimiento celular desordenado (tumor) y una colonización tisular (metástasis), todo ello determinado por una mutación inicial seguida de la acumulación de otras mutaciones sucesivas. Cada cáncer es una situación distinta con peculiaridades dependientes del tipo de célula donde se origina, sus causas (etiología) y su mecanismo, el grado de malignidad y otros factores. Por consiguiente, se diferencia claramente de otras enfermedades genéticas, especialmente las monogénicas (tema 28) y las anomalías cromosómicas (tema 29) y se distingue de las enfermedades multifactoriales (pág. 399) por la necesidad en éstas de una confluencia de factores genéticos y ambientales para iniciar la enfermedad. Al igual que todas las enfermedades moleculares, el cáncer se manifiesta en alteraciones bioquímicas de todo tipo a nivel celular, tisular y, particularmente, por la aparición de marcadores tumorales y tisulares (pág. 434).

Es, posiblemente, la enfermedad estudiada más intensamente y desde más puntos de vista diferentes, habiendo pasado en los últimos años de un conocimiento casi nulo a un gran avance en muchos aspectos moleculares, desde aquéllos que explican la proliferación excesiva y la pérdida de células por apoptosis (muerte celular), hasta los mecanismos por los que el cáncer genera las metástasis. Nuestro objetivo es presentar una visión global e integrada de todos esos aspectos, que ya han beneficiado a la Biología Molecular básica y son necesarios para comprender las aplicaciones, presentes y futuras, en especial para configurar una nueva capacidad clínica en cuanto a pronóstico, diagnóstico y tratamiento. La disponibilidad actual de fármacos para el tratamiento del cáncer por quimioterapia se verá superada algún día por novedosos tipos de abordaje basados en la terapia génica, así como en la inhibición específica de la angiogénesis, la metástasis y la apoptosis.

30.2 EL CÁNCER COMO ENFERMEDAD GENÉTICA

La participación de los genes en la carcinogénesis puede seguir dos tipos de mecanismo.

- Por un lado, algunos cánceres se deben a **cambios epigenéticos**, aún no bien conocidos, que afectan a la expresión de algunos genes y pueden permanecer en estado latente durante toda la vida de la célula. Por ejemplo, son abundantes los agentes mutágenos que modifican la expresión génica alterando la metilación del DNA (pág. 273).
- Sin embargo, la mayoría de los cánceres se deben a cambios genéticos inducidos por mutaciones. La célula inicial afectada crece con mayor rapidez de lo normal y, si no existe un control de ese crecimiento, llega a formar un **tumor primario** en el tejido circundante (**neoplasia**) o un **tumor secundario (metástasis)** en tejidos lejanos.

Las mutaciones del genoma responsables de la enfermedad son producidas por agentes cancerígenos, de muy variadas naturaleza y características (pág. 354), así como por errores espontáneos o inducidos en la replicación y reparación del DNA; por ello, el análisis de la carcinogénesis se hace a la luz del efecto mutagénico. Hoy son bien conocidas como causa de cáncer las mutaciones que afectan esencialmente a genes de dos tipos: los **protooncogenes** y los **genes oncosupresores** (o **genes supresores de tumores**). Esto justifica una de las características del cáncer, el ser una **enfermedad poligénica**, lo que no debe confundirse con su carácter monoclonal (pág. 413). Como resultado de la mutación, los genes expresan proteínas alteradas, sobreexpresan las normales o no expresan éstas, según los casos; cualquiera de estos efectos modifica funciones de la célula tan básicas como el control de su división (pág. 422), lo que se manifiesta en una proliferación excesiva y continua o en una reducción de la muerte celular programada o **apoptosis** (pág. 431).

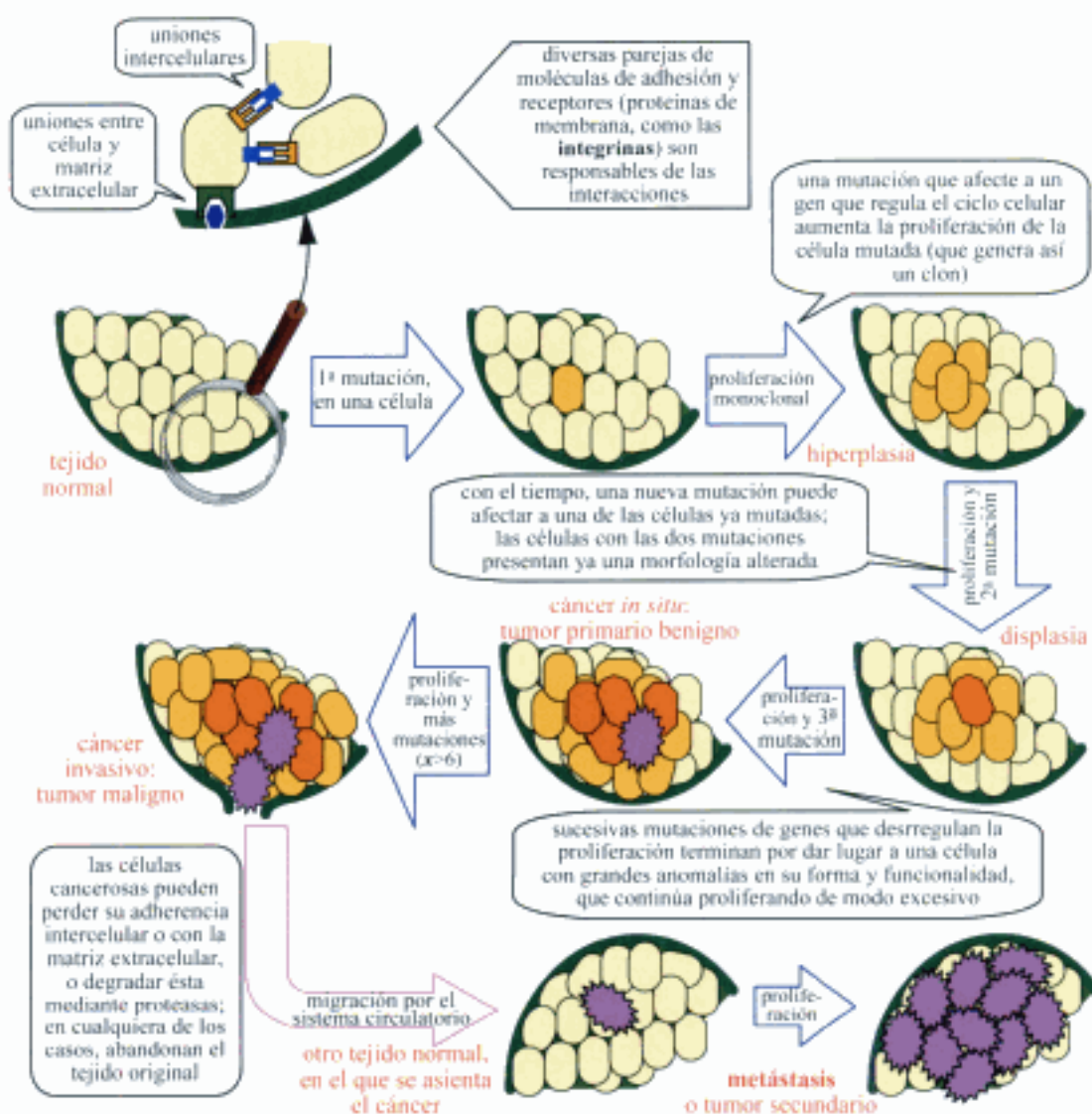


30.3 ETAPAS CARACTERÍSTICAS EN EL DESARROLLO DEL CÁNCER

El cáncer, una vez originado, se manifiesta de forma variable según las circunstancias, el tipo de cáncer, la edad del individuo, etc. A grandes rasgos, su desarrollo suele agruparse en tres fases, con posibles subetapas de difícil delimitación: la formación del tumor primario benigno, su progresión *in situ* y la invasión y aparición de metástasis o tumor maligno.

30.3.1 Tumor primario benigno

Esta primera etapa del desarrollo del tumor, llamada **tumorigénesis**, **carcinogénesis tumoral** o **transformación neoplásica**, es la más estudiada y mejor conocida. Se inicia con la llamada "célula progenitora del cáncer", que surge en un determinado tejido por una mutación en algún gen (pág. 412). Su proliferación da lugar a un clon (pág. 187) de células mutadas, un **clon neoplásico**. Este comienzo **monoclonal** es la primera característica del cáncer, que lo diferencia de las enfermedades multifactoriales (pág. 399); a pesar de ese origen común, las células del clon inicial se hacen genéticamente distintas por el acúmulo de nuevas mutaciones. Así, la mutación iniciadora confiere a la célula la **susceptibilidad tumoral**, una predisposición selectiva para la proliferación desmedida. La acumulación en las células hijas de sucesivas mutaciones en genes implicados en el control del ciclo celular o de la apoptosis, que favorezcan la proliferación o la inmortalidad, conduce progresivamente a las propiedades genéticas anormales del cáncer.



Cuantitativamente, se cree que son necesarias cerca de 6 mutaciones para que una célula normal se convierta en cancerosa. Dada la baja tasa de mutación (del orden de 10^{-6} por gen y célula), resulta prácticamente imposible que dicha conversión tenga lugar con mutaciones independientes. La probabilidad de que una de las cerca de 10^{14} células del cuerpo humano sufra 6 mutaciones es mínima ($10^{14} \times 10^{-6} \times 6 = 10^{14} \times 10^{-36} = 10^{-22}$). Es preciso, pues, que exista un mecanismo de multiplicación selectiva con cada nueva mutación.

30.3.2 Cáncer *in situ*

Corresponde a la situación en que las células del tumor primario no han escapado del tejido donde se originaron. Puesto que el tumor carece de capacidad invasiva, el pronóstico clínico es en general benigno y con frecuencia puede corregirse por simple extirpación quirúrgica. Sólo resulta grave si por su localización afecta a vasos sanguíneos, músculos o nervios críticos, en cuyo caso, aun sin ser formalmente un cáncer, podría considerarse como una neoplasia maligna.

30.3.3 Tumor maligno o cáncer propiamente dicho

La etapa que define y caracteriza el cáncer, responsable de la conversión de la célula neoplásica en cancerosa, es su progresión o propagación tumoral. A pesar de su trascendencia, ha sido la menos estudiada; sólo en los últimos años se le está prestando mayor atención y se conocen en parte sus aspectos moleculares. Puede dividirse en **varias subetapas**, de difícil subordinación.

30.3.3.1 Escape celular o invasividad

Las células constituyentes de tejidos normales se mantienen doblemente adheridas entre sí gracias a uniones intercelulares y a uniones de las células con el **espacio o matriz extracelular**. Ésta es una estructura formada por fibrillas de proteínas fibrosas (colágeno y elastina), secretadas por las propias células, que actúan a modo de soporte para empaquetar o separar las células o para proporcionar elasticidad o fuerza de sostén. En estas interacciones intervienen proteínas transmembrana, llamadas **integrinas**, de varios tipos, siendo la mejor descrita la fibronectina. Por su dominio intracelular, las integrinas se unen a los filamentos de actina del citoesqueleto de cada célula, mientras que a través de su dominio extracelular tienen afinidad por proteínas receptoras de las células vecinas (adhesión célula-célula) y por proteínas fibrosas de la matriz extracelular (adhesión célula-matriz).

Ambos tipos de adhesión se ven alterados en el cáncer y, como consecuencia, la célula puede desprenderse del tejido, convirtiéndose en una célula invasora de otros tejidos. La liberación de las células es consecuencia también de la acción de enzimas proteolíticas o **proteasas**, específicas, que destruyen la matriz extracelular, permitiendo que sea atravesada por la célula despegada de sus vecinas. Entre ellas, son destacables las llamadas **metaloproteasas de matriz extracelular** (MMP), que degradan casi todos los componentes de la matriz extracelular, por lo que intervienen en procesos fisiológicos de remodelación tisular, tales como desarrollo embrionario, crecimiento óseo y cicatrización de heridas. La actividad proteasa aumenta especialmente en las neoplasias malignas, posiblemente como consecuencia de un aumento de la biosíntesis de estas enzimas. En consecuencia, la inhibición de las proteasas adquiere importancia con fines terapéuticos, para el tratamiento del cáncer. Algunos de estos inhibidores actúan sobre la activación de las proteasas a partir de sus precursores inactivos (proenzimas o zimógenos, pág. 336), mientras que otros bloquean directamente la acción de la proteasa activa.

30.3.3.2 Vascularización o angiogénesis

Consiste este proceso en la capacidad de un tejido para formar una red vascular propia mediante el desarrollo de nuevos vasos sanguíneos, por mecanismos cuyos detalles se conocen aún de forma incompleta. Las células tumorales producen agentes angiogénicos que actúan sobre células endoteliales vecinas para dar lugar a la formación de una nueva red de venas, arterias y capilares. Ello permite que el tumor crezca mucho y muy rápidamente, al asegurar la llegada de oxígeno y nutrientes; la vascularización puede asimismo favorecer la migración de algunas células cancerosas al torrente circulatorio para difundirse a otros tejidos, donde generará una metástasis. La angiogénesis es uno de los objetivos de los tratamientos anticancerosos actuales, haciendo uso de la existencia de proteínas que inhiben el proceso, como endostatina y angiostatina.

30.3.3.3 Formación del tumor secundario o metástasis

Mediante la invasión o diseminación celular, las células del tumor primario pueden fijarse en tejidos normales, para multiplicarse y diferenciarse en ellos, dando lugar al llamado **tumor secundario o metástasis**. Esta etapa es de enorme gravedad clínica y constituye la nota distintiva letal, irreversible, del tumor maligno, porque supone la destrucción del tejido sano que rodea a las células tumorales y la consiguiente dificultad de eliminación por intervención quirúrgica.

30.4 MECANISMOS DE TRANSFORMACIÓN DE CÉLULA NORMAL EN TUMORAL

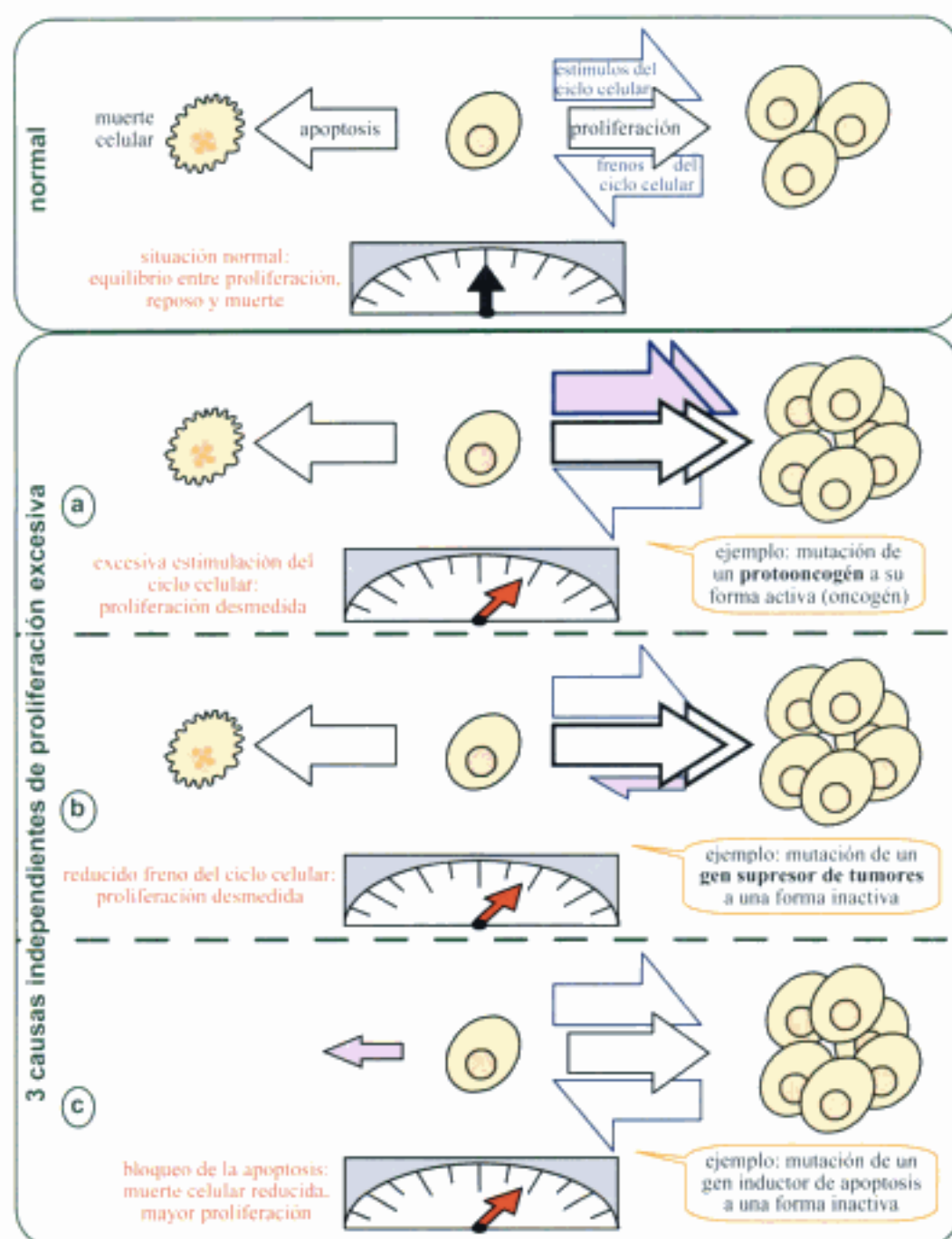
30.4.1 Equilibrio entre proliferación y muerte celular

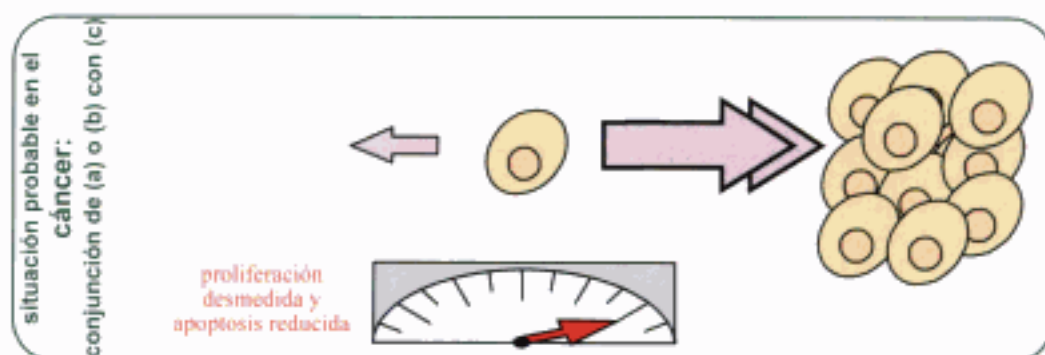
La vida media de la célula, o período de tiempo transcurrido desde su formación hasta su destrucción, es muy variable (horas, días, meses, etc.). Las células mantienen controlada esta capacidad intrínseca de crecimiento tanto en la embriogénesis y el desarrollo como a lo largo de la vida del individuo adulto. Algunas, como las precursoras hematopoyéticas o las epiteliales, con vida media muy corta, experimentan una proliferación constante para dar lugar a

otras muchas células; otras, como las neuronas, dejan de dividirse una vez diferenciadas. Salvo éstas, todas las células se renuevan continuamente en un proceso de "recambio celular". En condiciones fisiológicas, se requiere una estricta regulación genética para mantener esta homeostasis o equilibrio celular de forma adecuada a las necesidades de cada tejido, mediante la contraposición de dos procesos:

- La formación por mitosis de nuevas células (en el ciclo de división celular, tema 9) determina la tasa de **proliferación**. Se asocia a un proceso simultáneo de diferenciación celular, particular para cada tejido o estado del desarrollo (aspecto que no se considera en el presente libro). El control de la proliferación comprende dos grandes aspectos, relacionados pero generalmente no considerados de la forma global necesaria para facilitar su comprensión: las señales extracelulares (pág. 420) y la regulación del ciclo celular (pág. 422).
- La **apoptosis**, muerte celular programada o suicidio celular, es un mecanismo fisiológico de eliminación de células al final de su vida activa, como parte del recambio celular necesario para contrarrestar la formación continua por división celular. Igualmente, interviene en la formación de tejidos (por ejemplo, células del cristalino y de la piel), en la morfogénesis (por ejemplo, formación de los dedos) y en la eliminación de células con daño genético. Los detalles de la apoptosis se analizan posteriormente (pág. 431).

La alteración del equilibrio en condiciones patológicas (neoplasias) da lugar a un crecimiento incontrolado, que puede deberse a una excesiva proliferación o a una reducida muerte celular. Parece que normalmente el cáncer se origina por una conjunción de ambas causas.





30.4.2 Genes responsables del cáncer

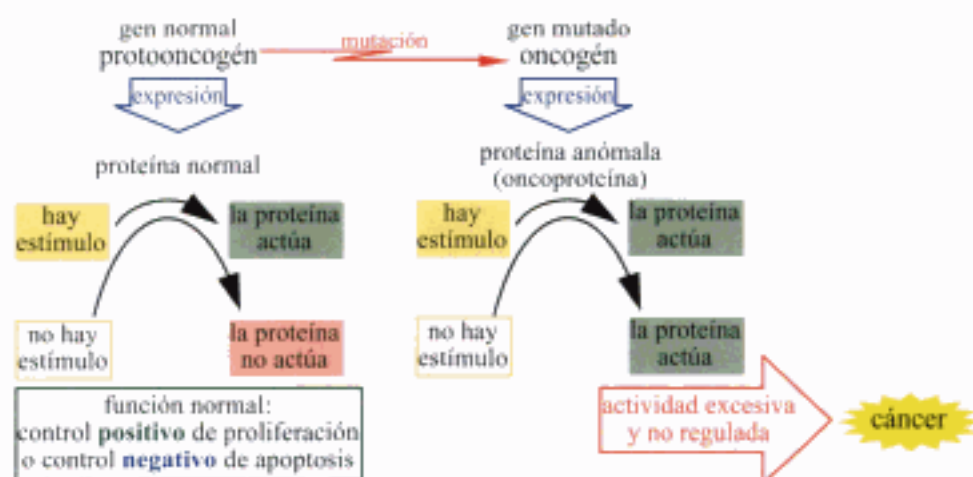
Las mutaciones, presentes en las células cancerosas, que provocan la aparición de las características tumorales afectan a genes cuyos productos proteicos regulan la proliferación celular, la apoptosis o la senescencia celular (envejecimiento). La proliferación y la apoptosis se estudiarán progresivamente en este tema, describiendo tanto los mecanismos normales de regulación como sus alteraciones conducentes al cáncer.

Primeramente, debemos definir la existencia de dos tipos de genes cuya alteración se relaciona con el proceso canceroso: oncogenes y genes oncosupresores (y sus productos respectivos, oncoproteínas y proteínas oncosupresoras).

30.4.2.1 Mutación de protooncogenes: oncogenes

a) Concepto

Se da el nombre de **oncogén** a la forma mutada de un gen normal o **protooncogén** que causa cáncer como consecuencia de una **ganancia de función** de la proteína expresada por el gen. El protooncogén codifica una proteína normal, generalmente relacionada con la proliferación celular o la apoptosis, que actúa sólo cuando recibe señales reguladoras específicas. Por el contrario, la forma mutada u **oncogén** expresa una proteína anormal (**oncoproteína**) que se mantiene activa independientemente de las señales reguladoras, por lo que el resultado es una ganancia de función de la proteína. Se puede decir, en consecuencia, que los oncogenes son el resultado de una mutación activadora de la función normal de los protooncogenes. Ésta conduce, según los casos, a una proliferación descontrolada o a una apoptosis reprimida y, en definitiva, a la aparición del cáncer.



b) Tipos de protooncogenes

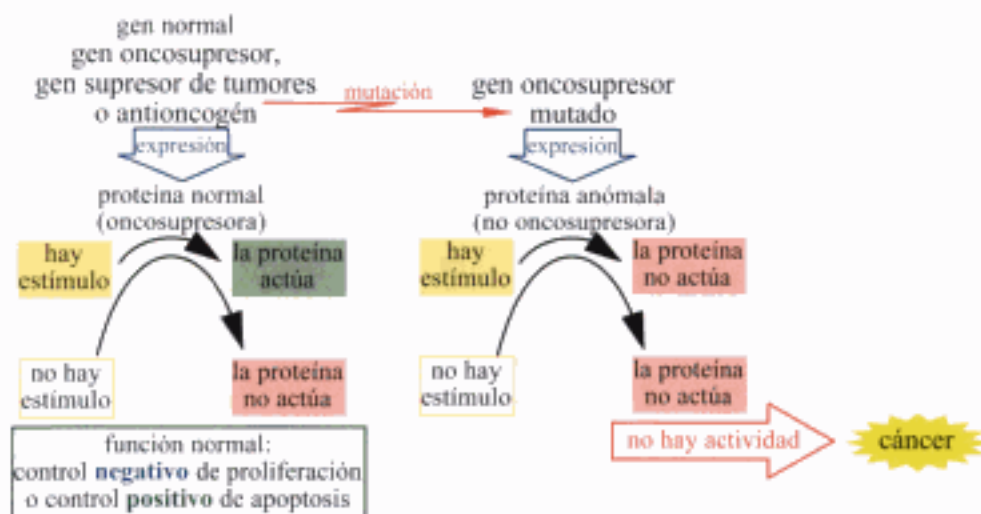
Se han identificado protooncogenes que codifican proteínas muy diversas:

- **Factores estimuladores del crecimiento celular.** La mutación oncogénica origina una producción excesiva del factor proteico o una mayor actividad de éste.
 - Ejemplo: el protooncogén *sis* codifica la cadena B del factor de crecimiento derivado de plaquetas (PDGF).

30.4.2.2 Mutación de genes oncosupresores

a) Concepto

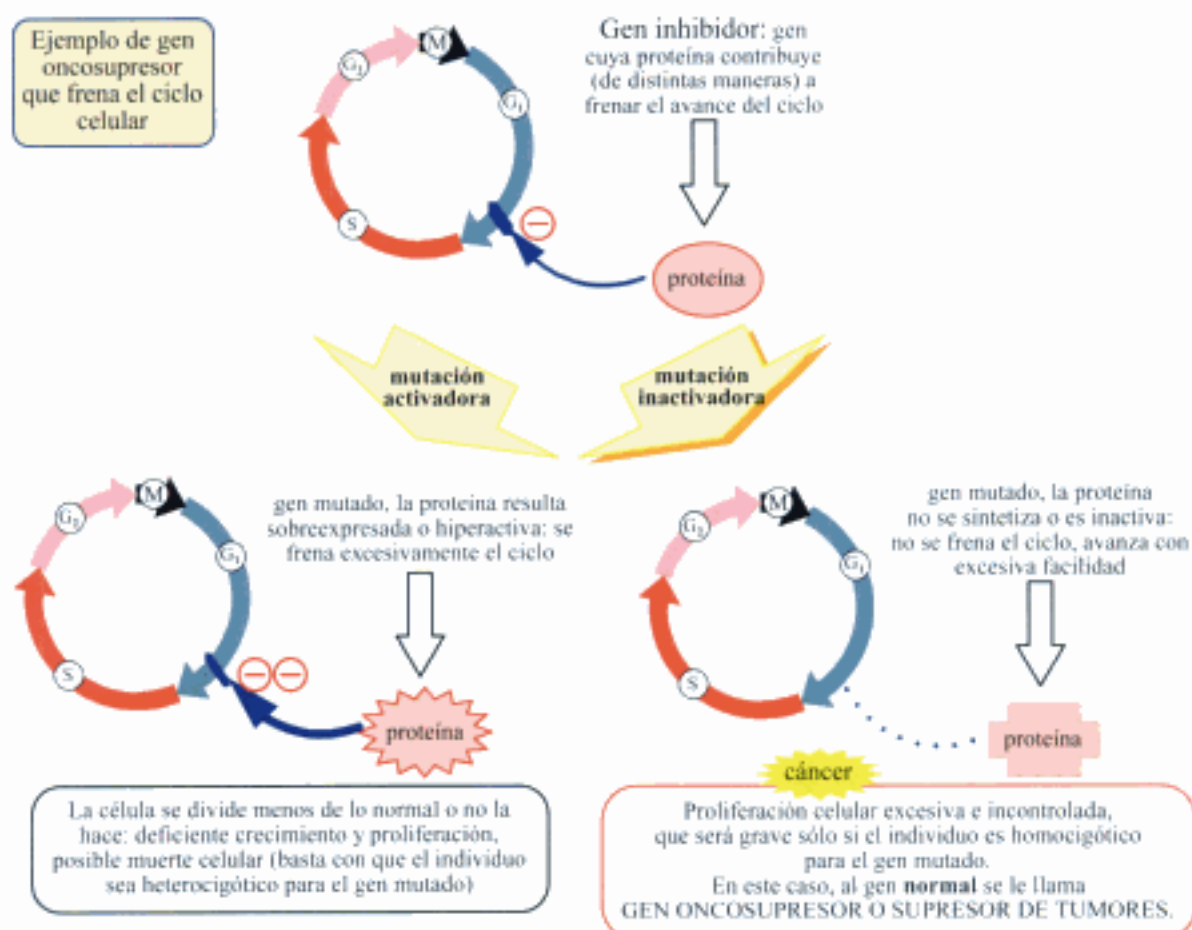
Al igual que los protooncogenes, se trata de genes que causan cáncer como consecuencia de una mutación experimentada por la forma normal, llamada **gen oncosupresor**, **gen supresor de tumores** o **antioncogén** (porque su acción es contraria a la de los oncogenes, aunque la ejercen por un mecanismo distinto). La forma normal del gen (**gen oncosupresor**) codifica una proteína normal (**proteína oncosupresora** o **antioncogénica**) que actúa deteniendo la proliferación o bien induciendo la apoptosis. Cuando una mutación del gen (gen oncosupresor mutado) conduce a la síntesis de una proteína no funcional o impide esa síntesis ("pérdida de función"), la proliferación deja de estar controlada o la apoptosis nunca tiene lugar, por lo que aparece una proliferación excesiva, a menudo con acumulación de daños genéticos en las células. Dicho de otro modo, la **versión mutada** de un gen oncosupresor pierde su función normal, reguladora de la proliferación, y da lugar al cáncer.



b) Tipos de genes oncosupresores

Al igual que los protooncogenes, los genes oncosupresores codifican proteínas que actúan en distintos puntos de las rutas de señalización, del control del ciclo celular y de la apoptosis:

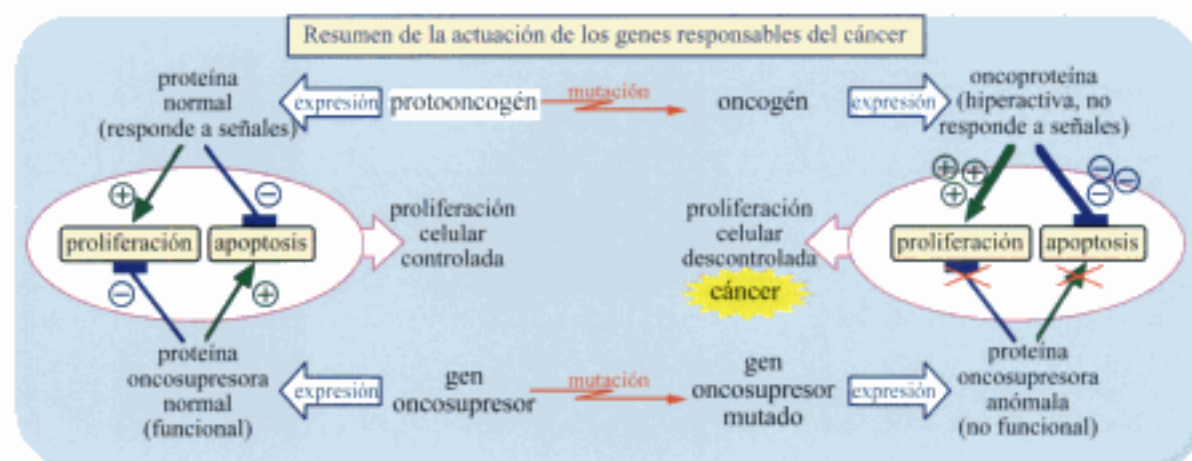
- **Factores inhibidores del crecimiento celular.** La mutación del gen oncosupresor correspondiente anula la funcionalidad de la proteína sintetizada.
 - Ejemplo: el gen oncosupresor *dcc* codifica una proteína de adhesión celular, ausente en el carcinoma de colon.
- **Receptores de esos factores inhibidores o de hormonas que frenan el crecimiento celular.** La forma mutada del gen oncosupresor codifica un receptor insensible a su ligando o que no transmite la señal al interior de la célula.
 - Ejemplo: el gen oncosupresor que codifica el receptor de membrana para el factor β de crecimiento transformante (TGF- β).
- **Proteínas citoplasmáticas que intervienen en los sistemas de transducción de señales** acoplados a los receptores anteriores. La situación varía según sea la función concreta de la proteína.
 - Ejemplo: el gen oncosupresor *nf1* codifica la neurofibromina, proteína que estimula la actividad GTPasa de Ras, en consecuencia inactivando ésta (pág. 422). En la enfermedad de Recklinghausen, el gen *nf1* está mutado (pág. 352), por lo que la neurofibromina es defectuosa, la proteína Ras no libera el GTP y queda en forma activa más tiempo del normal, aumentando la proliferación.
- **Factores de transcripción** que dirijan la expresión de genes cuyos productos proteicos frenan el ciclo celular o producen apoptosis. La forma mutada del gen oncosupresor expresa una proteína no funcional.
 - Ejemplos: los genes oncosupresores *Rb* y *p53*; se estudian en detalle más adelante (págs. 430 y 431).
- **Proteínas que frenan el ciclo celular o producen apoptosis.** El gen mutado no expresa la proteína o da lugar a una forma no funcional.



c) Características de la expresión de los genes oncosupresores

A diferencia de los oncogenes, donde un solo alelo mutante afecta al fenotipo de la célula, la pérdida de la función de un gen oncosupresor sólo se produce cuando han mutado los dos alelos (individuos homocigóticos, genotipo oncosupresor mutado-oncosupresor mutado, o pp). Es decir, la forma mutada de un gen oncosupresor se expresa con carácter **recesivo** (págs. 96 y 395). Esto es así puesto que la mutación produce pérdida de función en la proteína, pero las moléculas de proteína sintetizadas a partir del alelo normal pueden suplir esa carencia, al menos parcialmente. Al igual que en el caso de los oncogenes, la herencia de los genes oncosupresores mutados es la correspondiente a su carácter recesivo, pero sólo tiene lugar cuando la mutación ha afectado a células germinales (pág. 434).

La probabilidad de mutación somática en los dos alelos de una misma célula es baja, por lo que la homocigosis generada *de novo* es muy infrecuente; sin embargo, los individuos heterocigóticos en un gen oncosupresor (que han heredado un alelo mutado) tienen predisposición a la aparición de cáncer, ya que sólo necesitan una mutación somática en el otro alelo para perder totalmente la función de la proteína oncosupresora.



Corresponde estudiar a continuación con detalle las tres rutas que regulan la proliferación celular y cuya alteración conduce a cáncer:

1. La transducción de señales extracelulares al interior celular hasta conectar con el ciclo celular.
2. La regulación del ciclo celular (proceso estudiado previamente en el tema 9).
3. La apoptosis o muerte celular programada.

30.5 CONTROL DE LA PROLIFERACIÓN POR SEÑALES EXTRACELULARES

La señalización iniciada por moléculas externas a la célula constituye, en general, el mecanismo de control de la proliferación celular, desencadenante de la regulación del ciclo celular por otras moléculas intracelulares. Los mecanismos por los que las señales extracelulares dan lugar a una respuesta en el interior de la célula se conocen genéricamente como **transducción** de la señal, constituyen un área de gran significado en biología celular e intervienen no sólo en el control de la proliferación celular, sino también en la respuesta fisiológica a una gran cantidad de sustancias, como parte esencial de la función de los distintos tipos celulares. Se estudian aquí, de forma sucinta, las señales procedentes del exterior que van a determinar las respuestas intracelulares que regulan en último extremo el ciclo celular y, con ello, la tasa de proliferación. Aunque se conocen también como **señales mitógenas**, debería reservarse este término para las sustancias que actúan estrictamente sobre la mitosis.

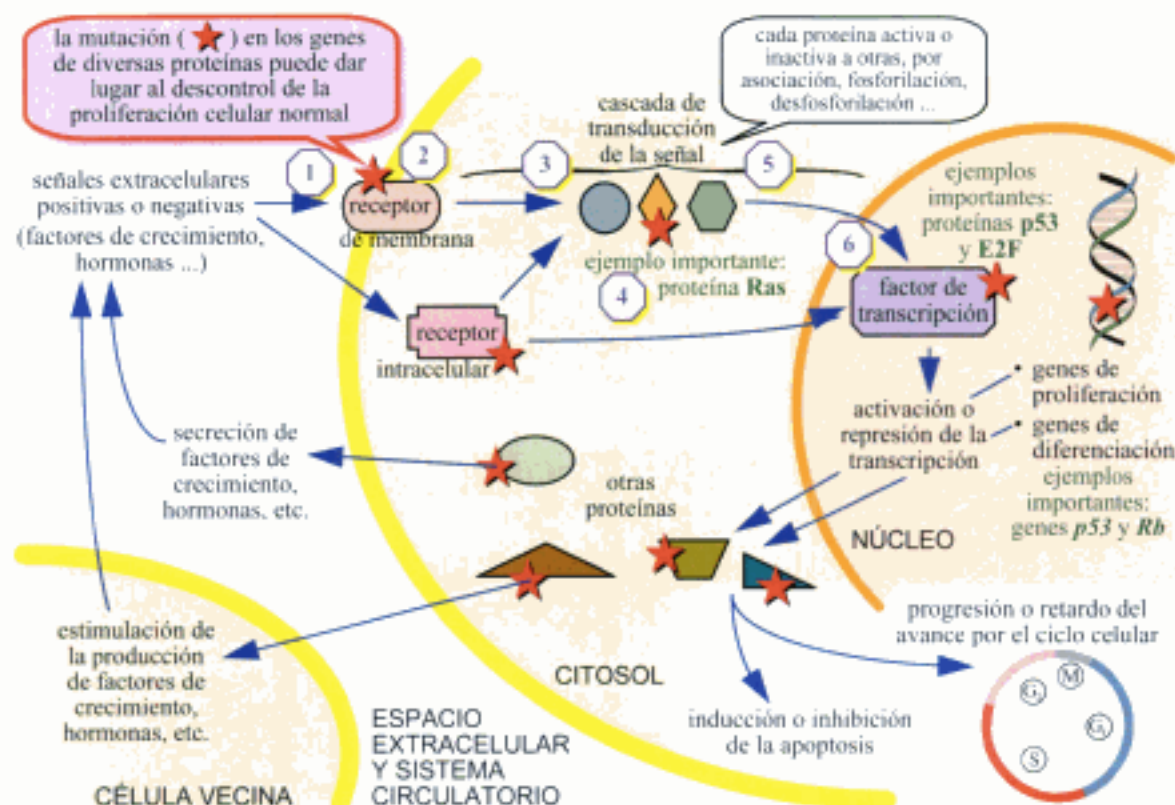
30.5.1 Tipos de señal

Las señales que favorecen la actividad del ciclo celular se conocen como **positivas** o **estimuladoras**. Existen también señales extracelulares **negativas** o **inhibidoras**, que actúan frenando la actividad del ciclo. En ambos casos, son muy variados los tipos y características de las moléculas que al incidir sobre su célula diana actúan como señal, a través de su interacción con receptores específicos.

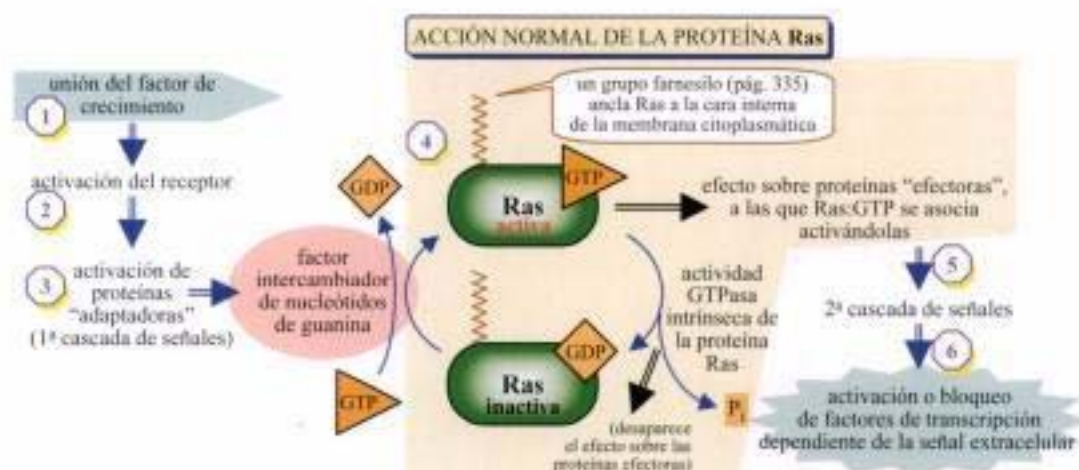
Tipos de señalización	
Por moléculas secretadas	
Señalización autocrina	La diana es la propia célula secretora
Señalización paracrina	La célula secretora y la célula diana son adyacentes
Señalización endocrina (u hormonal)	La célula secretora y la célula diana están alejadas; la molécula señal (hormona) se transporta por la sangre
Señalización intracrina	La molécula señal interacciona con un receptor intracelular en la propia célula
Señalización sináptica	La molécula señal (neurotransmisor) es secretada por una neurona e interacciona con el receptor en otra neurona o en una célula muscular que están en estrecha proximidad (sinapsis)
Por moléculas unidas a la membrana (señalización yuxtapuesta o yuxtacrina)	Interacción directa de moléculas de membrana con receptores de membrana de la célula diana
Por uniones comunicantes	Conexiones directas entre el citoplasma de las dos células
Moléculas señal	
Ejemplos	
Factores de crecimiento	Factor de crecimiento epidérmico (EGF) y factor de crecimiento derivado de plaquetas (PDGF)
Hormonas	
Péptidos y proteínas	Hormonas pancreáticas insulina y glucagón
Esteroides	Hormonas adrenales y sexuales
Derivados de aminoácidos	Hormonas tiroideas
Neurotransmisores	Sólo son relevantes en las neuronas
Otros	Nucleótidos, retinoides, derivados de ácidos grasos, óxido nítrico, monóxido de carbono ...
Tipos de receptor	
Intracelulares	Proteínas situadas en citosol y núcleo, que unen moléculas señal hidrofóbicas
De superficie celular	Proteínas integrales de membrana, que unen moléculas señal hidrofílicas
Con actividad enzimática intrínseca	Ejemplo: tirosina quinasas, como el receptor de PDGF y el de insulina
Asociados a canales iónicos	
Asociados a enzimas	
Asociados a proteínas G triméricas	Las proteínas G están acopladas, a su vez, a canales iónicos o a enzimas
Asociados a quinasas intracelulares	

30.5.2 Transducción de la señal extracelular al interior

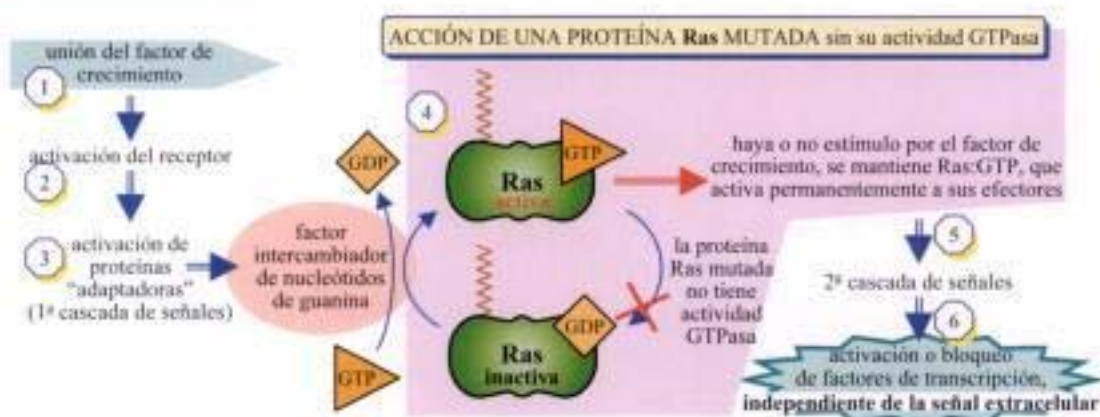
Aunque, como se ha indicado, existen muchas posibilidades de transmisión de la señal, por simplicidad la descripción se centrará en el mecanismo de activación en el que interviene la proteína **Ras**, causante de numerosos procesos cancerosos. Su activación es un paso esencial en la transducción de señales producidas por muchos factores de crecimiento.



1. La molécula señal actúa como **ligando** al interaccionar con gran afinidad y de modo reversible con el **dominio extracelular** de una **proteína receptora** específica, cuyo **dominio transmembrana** hidrofóbico está insertado en la bicapa lipídica de la membrana. El número de receptores de un tipo concreto en cada célula es muy variable, desde cientos a millones. En el caso de la ruta que activa a Ras, la señal procede de diversos receptores: para el factor de crecimiento derivado de plaquetas (PDGF), el factor de crecimiento epidérmico (EGF), la insulina, etc.
2. Al unirse el ligando al receptor, éste se estimula. En el caso del receptor de PDGF, forma un dímero y se activa una actividad proteína quinasa intrínseca, concretamente del tipo tirosina quinasa (págs. 327 y 425), fosforilando, con gasto de ATP, el **dominio intracelular** propio o bien otras proteínas.
3. El paso siguiente, que llamaremos "1ª cascada de transmisión de la señal", puede realizarse por varias vías en las que, en general, se producen modificaciones en la conformación de una proteína que a su vez permiten la interacción con otras proteínas "adaptadoras" o su fosforilación.
4. La propagación de la señal continúa mediante la activación de la proteína **Ras**, una proteína G monomérica (también llamada p21Ras, por su masa de 21 kDa). Las **proteínas G** actúan en numerosos procesos como "interruptores moleculares", al poder pasar de su forma inactiva, unida a GDP, al estado activo, unida a GTP (véanse como ejemplo algunos factores de traducción: eIF-2, pág. 310; eEF-1α y eEF-2, pág. 312; eRF, pág. 314). En este caso, el complejo activo Ras:GTP transmite la señal mediante su interacción con otras proteínas citoplasmáticas (proteínas "efectoras"), activándolas, con lo que comienza una "2ª cascada de transmisión de la señal". La consiguiente hidrólisis del GTP por Ras devuelve a ésta al estado inactivo, interrumpiendo así la transmisión de señal hasta que haya un nuevo estímulo. Debe añadirse que aunque Ras fue la primera descubierta y es la más importante, existen otras proteínas relacionadas que forman la familia Ras e intervienen en varios procesos celulares (como Rab, Rho y Arf).



Una **alteración** frecuente en una proporción elevada de tumores (especialmente de colon y páncreas) es la mutación del gen *ras*, dando lugar a una proteína Ras que es incapaz de hidrolizar el GTP, por lo que permanece continuamente en su forma activa Ras:GTP, desencadenando una señal de proliferación continua y no regulada, responsable de dichos cánceres.



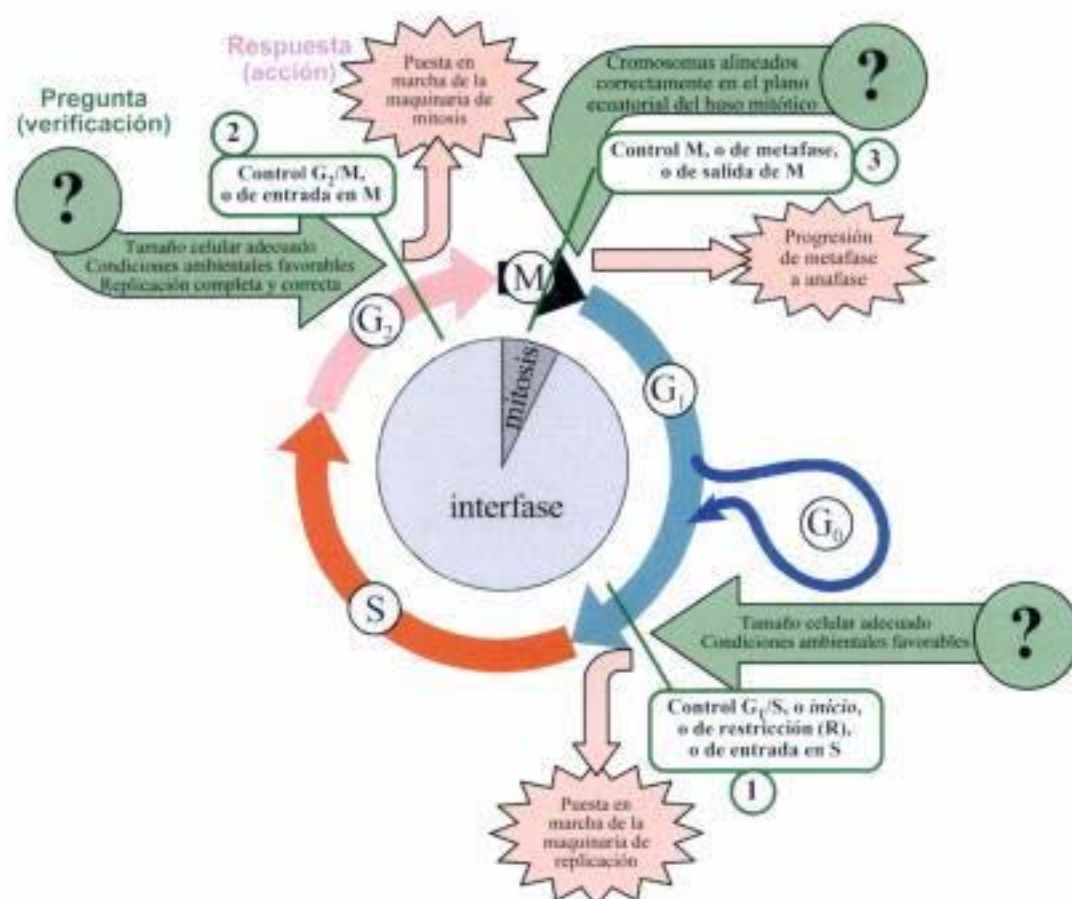
5. Cada proteína efectora activada por Ras produce a continuación cambios de conformación o fosforilaciones en sucesivas proteínas componentes de la "2ª cascada de señales", entre las que cabe citar Raf, MEK y MAP quinasas.
6. Finalmente, algunas de estas proteína quinasas de la 2ª cascada actúan sobre un factor de transcripción, al que fosforilan. Como consecuencia, éste cambia su conformación, lo que le permite unirse al DNA y activar o reprimir la transcripción de genes específicos que codifican las proteínas que regulan el ciclo celular y la apoptosis.

30.6 REGULACIÓN DEL CICLO CELULAR

Antes de describir los detalles de la regulación, debe indicarse que la repetición ordenada y controlada del ciclo celular define su duración característica, es decir, determina el grado de proliferación normal de las células en cada tejido. Cuando la proliferación celular es continua, excesiva y descontrolada se origina el cáncer. Como se ha indicado, esto puede ser consecuencia del acúmulo sucesivo (a veces durante años) de alteraciones o mutaciones en algunos de los genes anteriormente descritos (por ejemplo, *ras*). Analizaremos previamente los aspectos moleculares del proceso normal.

30.6.1 Puntos de control en el ciclo de división celular

La regulación del ciclo se ejerce en tres puntos bien definidos, que dependen en su conjunto de las condiciones ambientales más favorables (nutrientes, temperatura, sales) y la presencia de los factores de crecimiento celular necesarios. Los dos puntos de control principales se sitúan en la interfase, concretamente en las transiciones de G₁ a S y de G₂ a M. El tercer punto se localiza en plena mitosis. Gracias a este triple control, las células sólo inician una nueva fase de división cuando se han cumplido las condiciones necesarias en cada punto y se ha finalizado la fase anterior. Utilizando un símil, cada punto de control actúa a modo de "freno" de la etapa previa. Cualquier circunstancia anormal en uno de ellos hace cesar la división y la consiguiente proliferación celular.



1. El primero se denomina **punto de control en G₁/S**. Ocurre cerca del final de G₁ ($2n$, $2C$), antes de entrar en fase S, y una vez superado se dispara la replicación del DNA. Este punto se llama *inicio* en levaduras de gemación (como *Saccharomyces cerevisiae*) y *punto R* o *de restricción* en mamíferos, donde es el control principal del ciclo. Al ser la fase G₁ la más prolongada del ciclo (6-12 h), es en ella donde más intervienen las condiciones extracelulares (por ejemplo, del medio de cultivo), el único punto donde el ciclo responde a señales externas y progresa sin dependencia de estímulos mitóticos. El paso por este punto lo regulan las proteínas intracelulares llamadas **ciclinas de G₁** (pág. 424).
2. El segundo, o **punto de control en G₂/M**, ocurre al final de G₂ ($2n$, $4C$), antes del inicio de la mitosis M. Si la célula no supera este punto, permanece con ese complemento doble de dotación cromosómica. Su desregulación puede ser importante en la transformación neoplásica. El paso por este punto lo regulan las proteínas intracelulares llamadas **ciclinas de G₂ o mitóticas** (pág. 424).
3. El tercer control, **punto M**, se ejerce durante la mitosis ($2n$, $4C \rightarrow 2n$, $2C$), entre la metafase (cromosomas condensados dispuestos en el plano ecuatorial de la célula) y la anafase (separación de las cromátidas hermanas unidas al huso mitótico hacia cada polo celular). Asegura que la célula no se divida si hay errores en la formación del huso acromático o en la alineación de los cromosomas en la placa ecuatorial. Es el punto de control cuyo mecanismo es menos conocido y no se detallará en este texto, pero posiblemente depende también de las ciclinas mitóticas.

GLOSARIO particular para la regulación del ciclo celular:

Los **genes** se nombran con cursiva y minúsculas:

o todo en mayúsculas (dependiendo de la especie):

Las **proteínas** se nombran sin cursiva, con la primera letra en mayúscula:

cdc2

CDC2

Cdc2

(de *cell division cycle*,
ciclo de división celular).

Cdk = *cyclin-dependent kinase*, quinasa dependiente de ciclina. Se trata de una familia de proteínas con actividad proteína-quinasa sobre residuos de serina o treonina de determinadas proteínas sustrato. Sólo son activas como quinasas cuando están asociadas con una ciclina.

Ciclina = miembro de una familia de proteínas que se unen a las proteínas Cdk activándolas. Su nombre procede de las variaciones cíclicas que experimenta su concentración en la célula, en paralelo con el ciclo de división. Se clasifican en **ciclinas de G₁**, o de inicio y **ciclinas de G₂ o mitóticas**.

30.6.2 Proteínas que controlan el ciclo celular

En los puntos de control G₁/M y G₂/S intervienen dos tipos característicos de proteínas, las **ciclinas** y las **proteína quinasas**, que se asocian formando complejos Cdk-ciclina para controlar la actividad del ciclo celular. Se estudiarán sucesivamente los dos tipos de proteínas y la activación del complejo formado por ambas, para poder abordar después la regulación del ciclo celular por los complejos Cdk-ciclina.

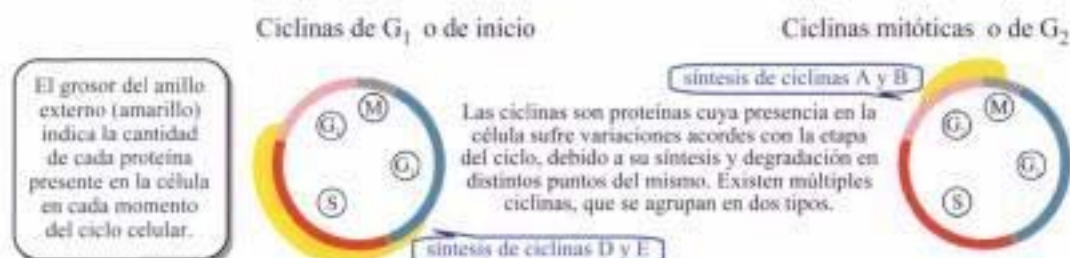
30.6.2.1 Ciclinas

El nombre de este grupo de proteínas se debe a que su concentración en la célula fluctúa de acuerdo con la etapa del ciclo celular, de modo que están presentes en una etapa concreta y desaparecen durante el resto del ciclo.

Existen numerosas ciclinas; las principales en humanos son las ciclinas A, B, D y E. Algunas de ellas están relacionadas entre sí; por ejemplo, las ciclinas B1 y B2, o las D1, D2, D3 y D4. La función exacta de otras ciclinas (C, F, G, H e I) se conoce peor; la ciclina F parece estar implicada en el control de la transición entre las fases S y G₂.

a) Ciclinas de la fase G₁ o ciclinas de inicio

Son aquellas que intervienen regulando el punto de control G₁/S. En humanos pertenecen a este grupo las ciclinas del tipo D (D1, D2, D3 y D4) y la ciclina E. Ambas se sintetizan durante la fase G₁ (por ejemplo, al entrar la célula en el ciclo desde G₀) y su concentración celular se regula por transcripción. Debido a su rápida degradación, tienen una vida media corta (unos 25 min para las D). El balance entre síntesis y degradación hace que aparezcan en la célula al final de G₁ y desaparezcan durante la fase S o al acabar ésta.

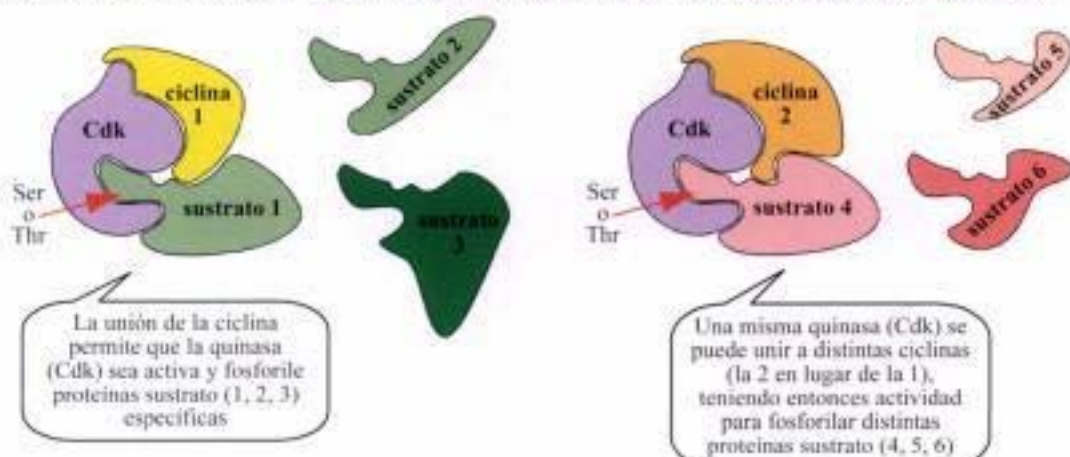


b) Ciclinas mitóticas o de la fase G₂

Este grupo, formado por las ciclinas A, B1 y B2, regula el punto de control G₂/M. En este caso se sintetizan durante las fases previas a la mitosis (S o G₂) y se degradan rápida y específicamente durante la mitosis. La ciclina A participa en el control de la replicación del DNA y en la entrada en mitosis. La ciclina B, que sólo participa en la mitosis, se considera la ciclina por antonomasia, pues es la homóloga en mamíferos de la primera ciclina descubierta al comenzar el estudio del control del ciclo celular en los oocitos de la rana *Xenopus laevis*, la que forma parte del factor promotor de la maduración o de la fase M (MFP en inglés).

30.6.2.2 Proteína quinasas dependientes de ciclinas (Cdk)

Constituyen una familia de proteínas con una gran homología de secuencia (superior al 50%), que fosforilan residuos serina y treonina de proteínas sustrato específicas. Se trata, por tanto, de serina/treonina quinasas (pág. 327). Sin embargo, sólo pueden ejercer esa actividad enzimática cuando están asociadas con una ciclina; por ello, se las llama quinasas dependientes de ciclinas, **Cdk** o **CDK** (de *cyclin-dependent kinases*). De acuerdo con ello, se considera que en el complejo activo (Cdk-ciclina) la Cdk es la subunidad catalítica y la ciclina la subunidad reguladora.



		Ciclinas de G ₂		Ciclinas de G ₁		Otras ciclinas menos conocidas				
		A	B	D	E	C	F	G	H	I
Proteína quinasas	Cdk1	G ₂ /M	G ₂ /M							
	Cdk2	S		(G ₁ /S)	G ₁ /S					
	Cdk3				G ₁ /S					
	Cdk4			G ₁ /S						
	Cdk5			(G ₁ /S)						
	Cdk6			G ₁ /S						
	Cdk7									✓
	Cdk8					✓				
	¿?							S/G ₂		

asociaciones de ciclinas de G₂ que controlan el ciclo en el punto indicado

asociaciones de ciclinas de G₁ que controlan el ciclo en el punto indicado

asociaciones de otras ciclinas, cuyo punto de control no se conoce

control por una Cdk desconocida

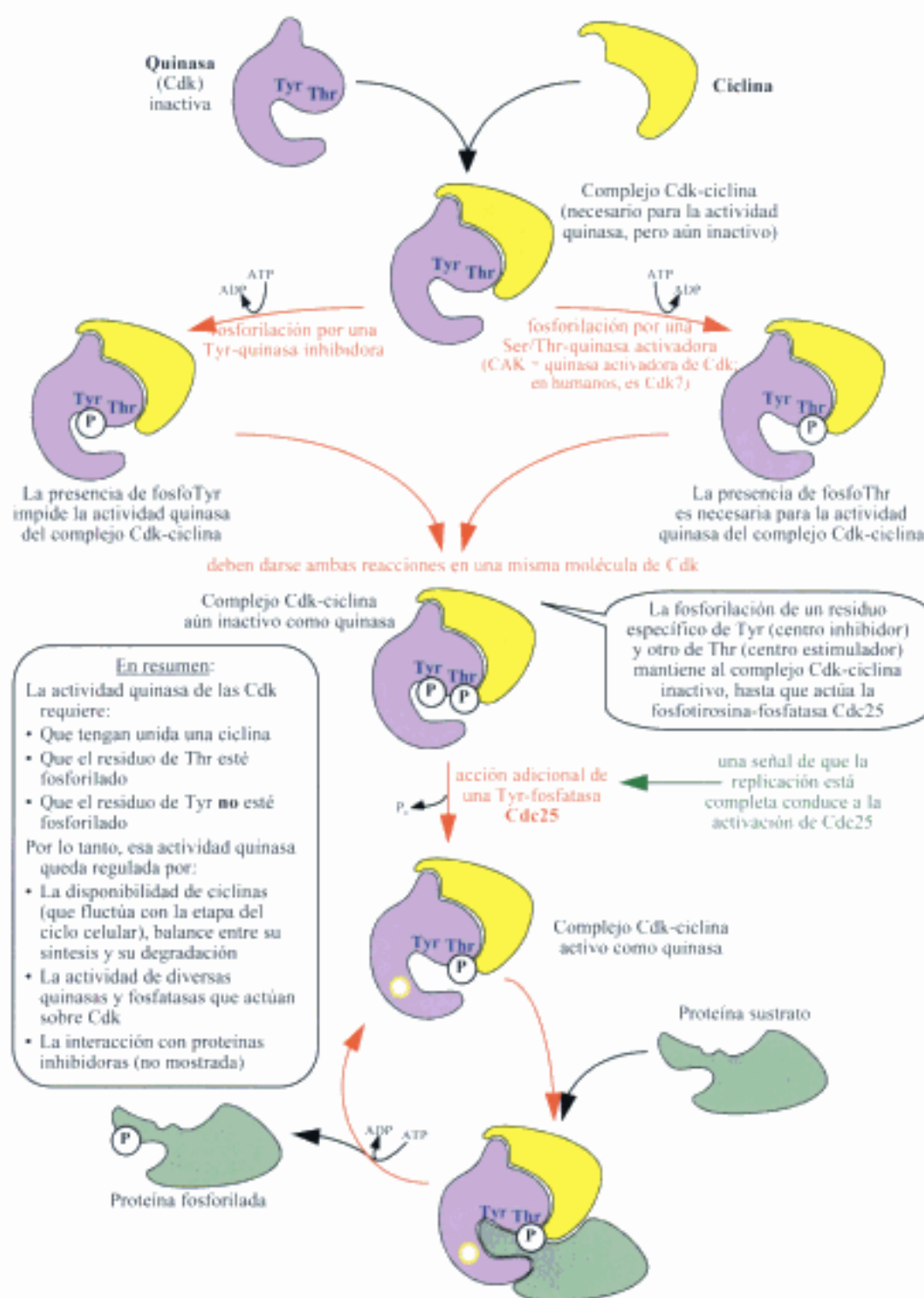
30.6.2.3 Actividad de los complejos Cdk-ciclina

La activación de las Cdk como proteína quinasas no consiste simplemente en su unión con una ciclina, sino que su regulación depende de otros mecanismos:

- La posibilidad de que se forme el complejo Cdk-ciclina depende de la disponibilidad de ambas proteínas; puesto que las Cdk están presentes de forma continua en la célula, son las ciclinas las que regulan.
 - Tasa de síntesis de las ciclinas:** como se ha indicado (pág. 424), la síntesis está controlada en función de la etapa del ciclo, aumentando en ciertos puntos del mismo, diferentes para las ciclinas A y B que para las D y E.
 - Tasa de degradación de las ciclinas:** determina, junto con la síntesis, la concentración disponible para su unión a las Cdk.
- Una vez formado el complejo Cdk-ciclina, su actividad como quinasa está sometida a regulación por otras proteínas.
 - Modificación de la Cdk por fosforilación y desfosforilación:** la actividad del complejo Cdk-ciclina precisa la fosforilación de ciertos residuos aminoácidos de la Cdk y no de otros, lo cual está regulado por distintas quinasas y fosfatasas, que a su vez están reguladas por la propia Cdk activa (véase figura en pág. 427).
 - Interacción de las quinasas con proteínas inhibidoras:** el último mecanismo regulador de la actividad de los complejos Cdk-ciclina se basa en la existencia de muchas proteínas que se unen a las Cdk, inhibiendo directamente su actividad; se las llama *inhibidores de los complejos Cdk-ciclina*, o **CKI**. Su efecto es, por tanto, bloquear el ciclo de división, por lo que funcionan como oncosupresoras (pág. 418).

Los CKI se agrupan en dos familias, en función de su estructura (existe similitud de sus secuencias), modo de acción y especificidad de sustrato:

familia	nombre de la proteína inhibidora	otros nombres		acción inhibidora
KIP	p21	CIP1	WAF1	Preferentemente sobre los complejos Cdk-ciclina de G ₁ y S. En menor medida, sobre complejos Cdk1-ciclina de G ₂
	p27	KIP1		
	p57	KIP2		
INK4	p16	INK4a	MTS1	Específicamente sobre los complejos de ciclinas de G ₁ con Cdk4 o Cdk6, pues compiten con las ciclinas D
	p15	INK4b	MTS2	
	p18	INK4c		
	p19	INK4d	ARF	

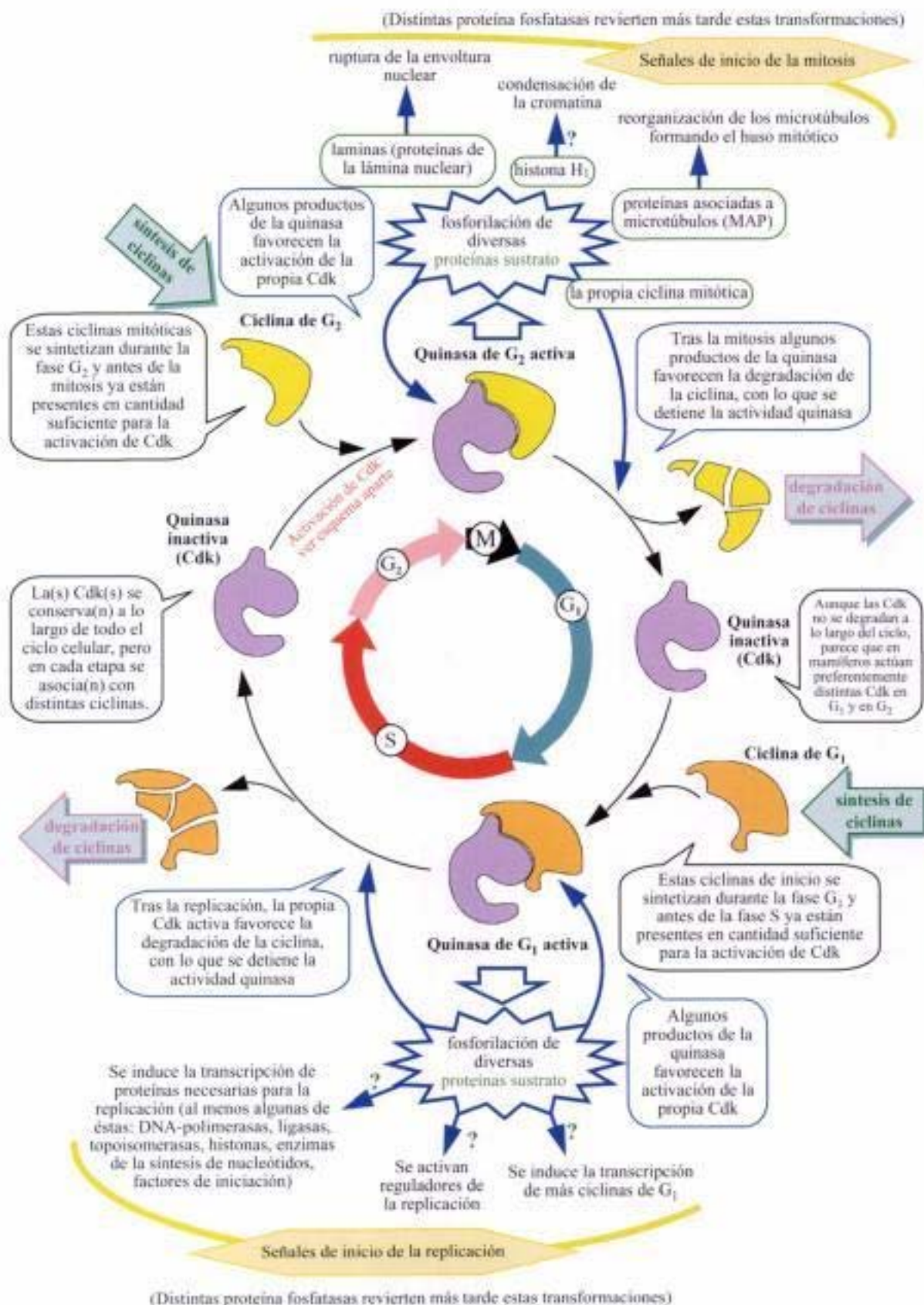


30.6.3 Regulación del ciclo celular por los complejos Cdk-ciclina y alteraciones en el cáncer

El descubrimiento de la relación entre distintos procesos neoplásicos y las alteraciones en algunos genes, especialmente aquellos con un papel clave en el control del ciclo celular, ha permitido comprender la trascendencia y los mecanismos de funcionamiento de este control. Distintas alteraciones en las proteínas que regulan el ciclo de división celular pueden tener como consecuencia la división celular excesiva o incontrolada, dando lugar a una proliferación celular desmedida con acumulación de anomalías genéticas y fisiológicas en las células, que pasan a constituir el tumor.

30.6.3.1 Estímulos de progresión por el ciclo celular y oncogenes que los alteran

Como ya se ha descrito, el avance por el ciclo celular se impulsa a grandes rasgos en dos puntos: la síntesis y degradación de ciclinas de G_1 , seguida de la activación de sus complejos Cdk-ciclina, y la etapa equivalente en G_2 . La activación de estos complejos conduce a respuestas variadas, mediadas por las proteína diana que resultan fosforiladas por la actividad quinasa. Se muestra a continuación una visión global de estos dos puntos de activación del ciclo.

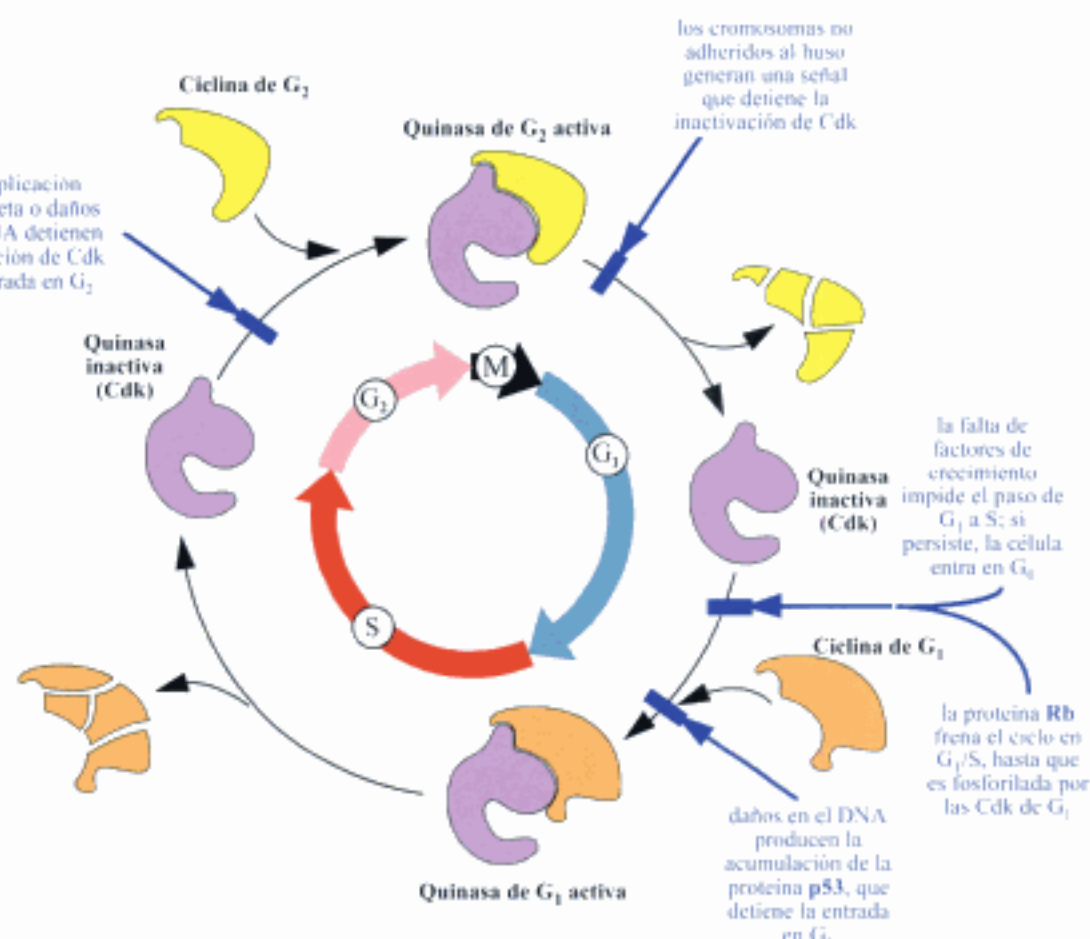


Esta estimulación normal de la progresión por el ciclo puede alterarse, conduciendo a un cáncer. La mutación de protooncogenes que codifican estas proteínas estimuladoras del ciclo los pueden transformar en oncogenes. Esto ocurre cuando las oncoproteínas respectivas actúan sin obedecer a la regulación, provocando una estimulación continuada del ciclo que conduce a la proliferación excesiva causante del cáncer.

Ejemplo de proteína reguladora del ciclo celular, codificada por un protooncogén	Efecto oncogénico resultante de la mutación
Ciclina	Síntesis excesiva Síntesis no regulada de acuerdo con el ciclo
Cdk	Síntesis excesiva Actividad quinasa excesiva o independiente de la unión de la ciclina
Proteína quinasa o proteína fosfatasa activadoras del complejo Cdk-ciclina	Síntesis o actividad excesivas
Proteína quinasa o proteína fosfatasa inhibidoras del complejo Cdk-ciclina	Proteína no funcional

30.6.3.2 Estimulos de detención del ciclo celular

La regulación normal del ciclo depende del balance entre los estímulos positivos que lo impulsan y los negativos o frenos que detienen el ciclo en los tres puntos de control, G₁/S, G₂/M y M (pág. 423), para asegurar la correcta preparación de la célula para la fase siguiente (por ejemplo, que se haya completado la replicación antes de que se condense el genoma y se divida la célula). Por ahora, los mecanismos de actuación de estos estímulos negativos se conocen peor a escala molecular que aquéllos por los que tiene lugar el control positivo; sin embargo, conceptualmente sus implicaciones en la regulación normal y alterada en el cáncer son similares en ambos casos.



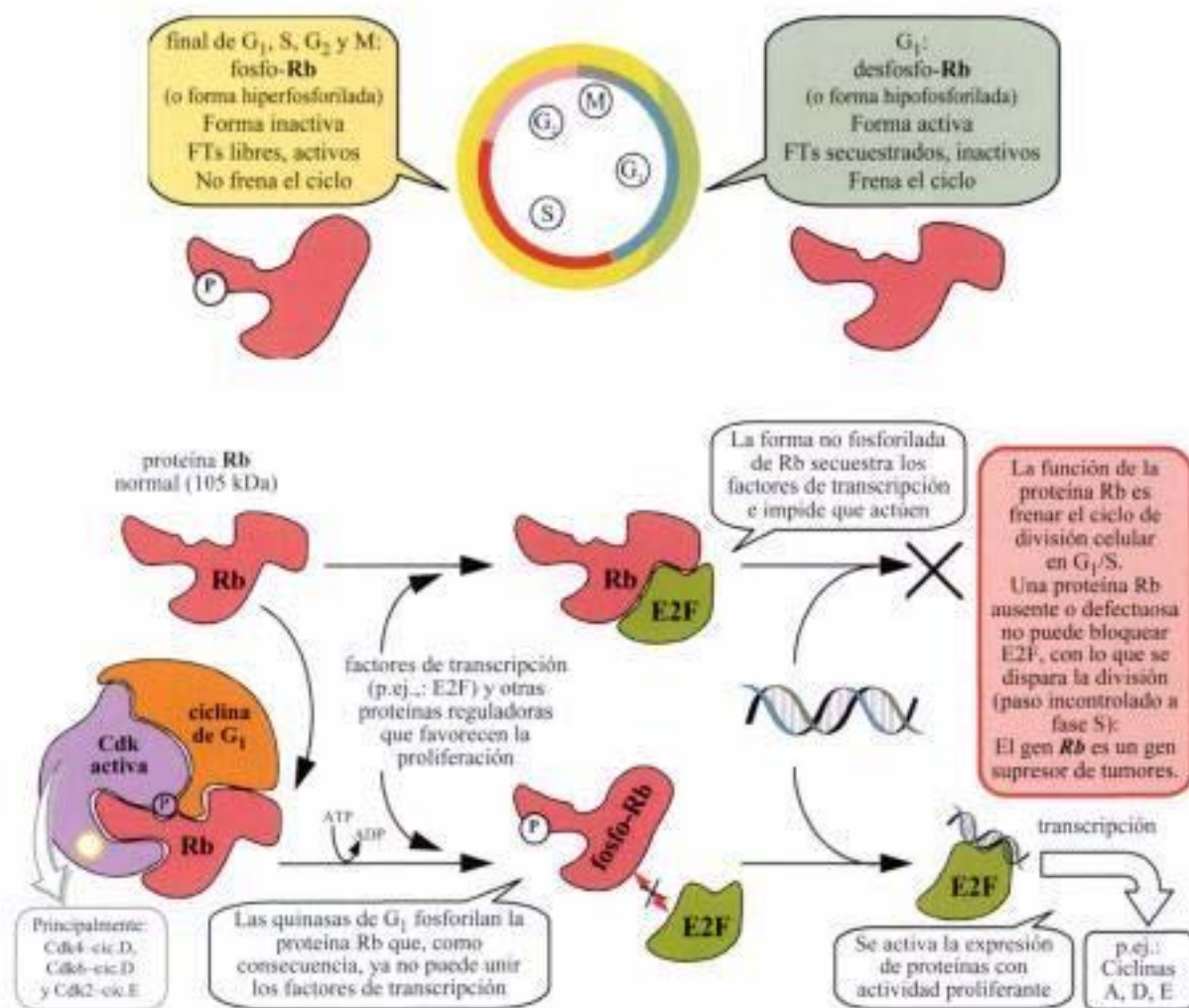
La aparición de cáncer por alteración del control negativo corresponde a genes oncosupresores, o supresores de tumores, cuya función normal es frenar el ciclo y, al verse mutados, dejan de hacerlo (pág. 418). Como consecuencia, la célula replica su DNA y se divide por mitosis sin atender a los puntos de control y, en consecuencia, prolifera

excesivamente y acumula defectos genéticos (por ejemplo, por una mitosis prematura antes de que se haya completado la replicación). Entre los múltiples genes supresores de tumores descritos, los mejor estudiados son *Rb* y *p53*, cuyos productos proteicos actúan sobre la regulación del ciclo. A continuación se detallan su función normal y su implicación en el cáncer.

30.6.3.3 El gen supresor del retinoblastoma (*Rb*)

El descubrimiento de una predisposición hereditaria para sufrir **retinoblastoma**, un tipo infrecuente de cáncer que afecta a los ojos, permitió identificar un gen (denominado *Rb*) de trascendencia en el control del ciclo celular. Su producto génico, la proteína del retinoblastoma (**Rb**, **pRb** o p105-Rb, por su masa de 105 kDa), se encuentra presente y actúa en el control de todo tipo de células normales; de ahí su importancia general, en contraste con el carácter restringido de actuación que pudiera deducirse de su nombre.

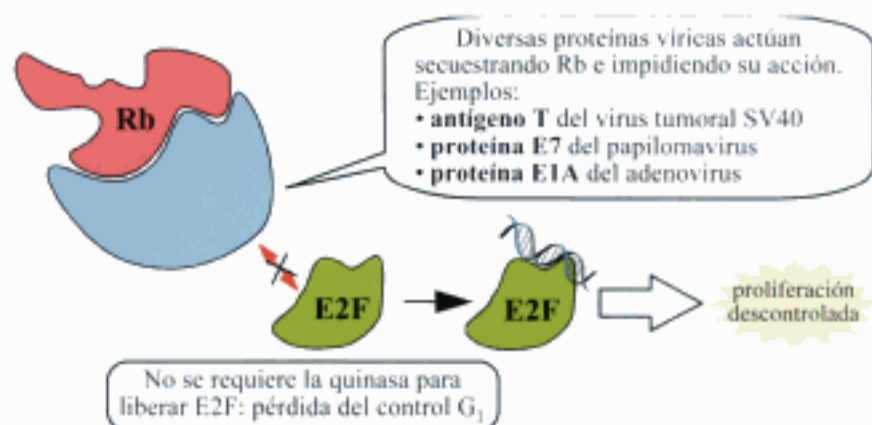
La proteína Rb se caracteriza por experimentar cambios de fosforilación a lo largo del ciclo celular, que modulan su actividad como freno en el punto de control G₁/S. Esta actividad se debe a la asociación de Rb no fosforilada con una gran cantidad de otras proteínas necesarias para la proliferación celular, en especial factores de transcripción inducibles (pág. 269) como **Myc** y **E2F**.



En una célula normal, la proteína Rb está siempre presente, pero su estado de fosforilación (y, por consiguiente, su actividad) varía a lo largo de las etapas del ciclo. Los factores estimuladores del crecimiento actúan como señal que favorece la activación de las proteínas Cdk, con lo que éstas pueden fosforilar Rb y anulan así su efecto inhibitor sobre la progresión del ciclo en G₁/S.

La pérdida o mutación de las dos copias (alelos) del gen *Rb* anula la función de la proteína Rb, suprimiendo el control por Cdk y provocando un paso prematuro a fase S. Aunque la inactivación del gen *Rb* es responsable de cánceres en diversos tejidos, afecta en especial a las células de la retina inmadura, conduciendo a la proliferación celular excesiva que causa el retinoblastoma. Los niños que heredan una copia defectuosa del gen *Rb* poseen un único alelo normal, por lo que es más probable que una mutación anule completamente la funcionalidad de Rb y tienen mayor susceptibilidad al desarrollo de tumores en la retina.

Finalmente, la acción de la proteína Rb normal se puede ver también impedida como consecuencia de ciertos virus, que tiene por ello acción promotora del cáncer independiente de la mutación del genoma humano.



30.6.3.4 El gen oncosupresor p53

Aproximadamente el 50% de los tumores humanos presentan mutaciones en un gen relacionado con el control del ciclo celular, denominado **p53** por codificar una proteína cuya masa molecular es de 53 kDa. Esto no quiere decir que la proteína **p53** anormal sea la única causa del 50% de los tumores, pues se acumulan mutaciones en otros genes cuyo efecto combinado es la enfermedad. Sin embargo, las lesiones del gen **p53** tienen un efecto claro sobre la progresión neoplásica.

El carácter oncosupresor del gen **p53** se debe a las funciones de la proteína normal sintetizada, que consisten en **inducir la apoptosis** y actuar como **freno en el punto de control G₁/S**. La concentración de la proteína **p53** en células normales es baja, por lo que sólo adquiere relevancia en situaciones especiales. Así, se ha comprobado un importante aumento de **p53** cuando existe daño en el DNA (por ejemplo, tras la exposición a radiaciones), cuando se altera la concentración de nucleótidos en la célula, en situaciones de hipoxia, bajo estímulos oncogénicos, etc. En esas condiciones, aumenta la concentración de la proteína **p53** y se prolonga su vida media (pasando de 5-10 minutos hasta 150 minutos). La presencia de daños en el DNA es detectada por varias proteínas, entre ellas una proteína quinasa que fosforila a **p53**. La forma fosforilada de **p53** es un factor de transcripción (pág. 264) que induce la expresión de algunos genes y reprime la de otros. Entre las proteínas cuya expresión aumenta destaca en primer lugar **p21**, que inhibe la actividad quinasa del complejo Cdk2-ciclina E, necesaria para superar G₁/S. En segundo lugar, **GADD** (*growth arrest on DNA damage*, detención del crecimiento por daños en el DNA), una proteína que, al igual que **p21**, secuestra el PCNA, factor de procesividad de la DNApol δ (pág. 149), impidiendo de este modo la replicación. En consecuencia, y de forma similar a Rb, la función normal de **p53** es frenar la progresión del ciclo, evitando que la célula replique ese DNA dañado. Más aún, es posible que la propia proteína **p53** participe o estimule la reparación del DNA.

Los defectos homocigóticos en **p53** conducen, por tanto, a una falta de control en el punto G₁/S y una reducción de la entrada en apoptosis, resultando una proliferación celular descontrolada que ocasiona tumores. En individuos heterocigóticos producen el síndrome de Li-Fraumeni, caracterizado por una predisposición al desarrollo de tumores. Por último, de manera análoga a lo que ocurre sobre Rb, diversas proteínas víricas generan tumores a través del bloqueo de **p53** (antígeno T de SV40, proteína E6 de papilomavirus y proteína E1B de adenovirus, por ejemplo).

30.7 APOPTOSIS, MUERTE CELULAR PROGRAMADA O SUICIDIO CELULAR

La apoptosis es un proceso fisiológico, controlado genéticamente, que se da en todo tipo de células de organismos multicelulares y por tanto en todos los órganos y sistemas, mediante el cual las células dañadas o no deseadas (que han cumplido su ciclo vital, pág. 415) activan mecanismos que conducen a su propia destrucción; de ahí el nombre alternativo de "suicidio celular". No debe confundirse con la necrosis, debida al efecto nocivo directo de agentes externos, de naturaleza mecánica (rotura), biológica (virus, bacterias), quimiofísica (drogas, toxinas, venenos), etc., que conducen a pérdida de la capacidad de regulación osmótica, hinchamiento, rotura de la membrana, salida del contenido intracelular y, finalmente, fagocitosis de los residuos celulares.

Se observa reducción de la apoptosis en condiciones fisiológicas caracterizadas por una gran proliferación celular (desarrollo embrionario y posnatal temprano, metamorfosis, mecanismo de acción de hormonas y factores paracrin, etc.), así como en situaciones patológicas tan diversas como el cáncer (la más representativa de una excesiva proliferación), enfermedades autoinmunes o infecciones víricas. De ahí la importancia del estudio del proceso apoptótico y su control y fallos en procesos patológicos.

El estudio de la apoptosis ha de hacerse en conexión con el de la proliferación celular. Como ya se ha indicado (pág. 415), la célula cancerosa se considera hoy día como una célula con una proliferación incontrolada debida probablemente a la suma de una activación de la proliferación y un bloqueo de la apoptosis. De hecho, la propia

quimioterapia del cáncer está imbuida de esta idea: el impacto más bien modesto de los tratamientos antiproliferativos, posiblemente causado por una baja capacidad de proliferación de muchos tumores, debe complementarse con tratamientos que induzcan la apoptosis.

30.7.1 Etapas y control del proceso de apoptosis

Desde un punto de vista celular y molecular, los cambios en la apoptosis parecen transcurrir en tres **etapas** relativamente definidas, denominadas inducción, ejecución y fase final.

30.7.1.1 Inducción

Se da este nombre a la reacción ante un estímulo o señal. Al igual que en la proliferación celular, es importante indicar que la inducción de la apoptosis requiere estímulos adecuados (señales positivas), tanto extracelulares como procedentes de la propia célula. Existen también señales negativas que actúan en sentido contrario, frenando la apoptosis.

30.7.1.2 Ejecución

Tras la inducción comienza el verdadero proceso de apoptosis, que consta de alteraciones diversas que conducirán a la muerte de la célula: cambios morfológicos y de la membrana con disminución de la adhesión a otras células y a la matriz extracelular, contracción del citoplasma, condensación de la cromatina, fragmentación del DNA, disminución del volumen nuclear seguida de su desintegración, rotura de mitocondrias con vertido de su contenido al citoplasma, aparición de protuberancias en la membrana celular (fenómeno conocido como **zeiosis**), etc. Estas alteraciones se deben a la activación de varias proteasas y nucleasas, que se convierten así en las principales moléculas responsables de la apoptosis.

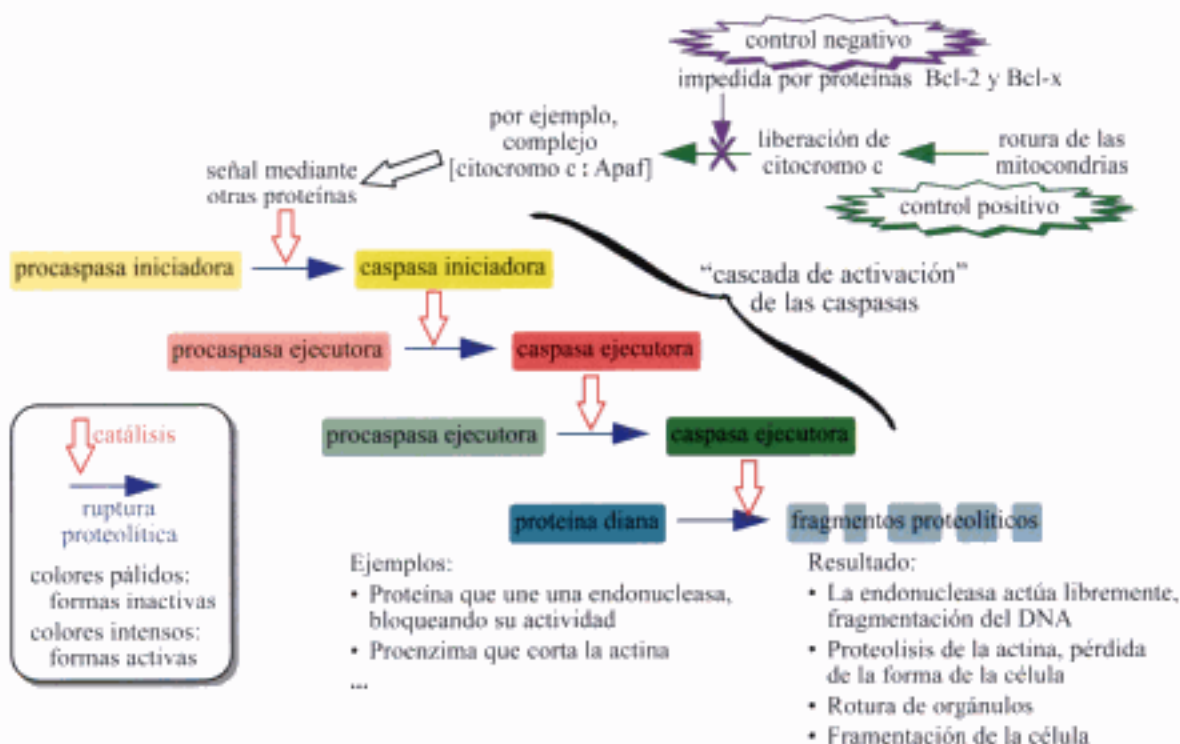
a) Proteasas: caspasas y su papel en el control de la apoptosis

Las **caspasas** son una familia de proteasas citosólicas (dentro de la clase EC 3.4.22, cisteína endopeptidasas). En humanos se han identificado 10, clasificadas en tres grupos de acuerdo con su especificidad de sustrato. Producen la proteólisis de proteínas diana o sustrato, entre las cuales cabe citar una proteína nuclear con actividad de reparación del DNA (PARP o poli(ADP-ribosa)polimerasa) y un factor de transcripción (SREBP). Las caspasas no sólo actúan degradando proteínas sustrato, sino que también activan a otras caspasas. Dependiendo de su posición en la cascada de activación, se distingue así entre caspasas "iniciadoras" y "ejecutoras".

son **cisteinproteasas**, es decir, tienen un residuo Cys implicado en el centro activo

caspasas

escinden la cadena polipeptídica de la proteína sustrato en posiciones posteriores a un residuo Asp



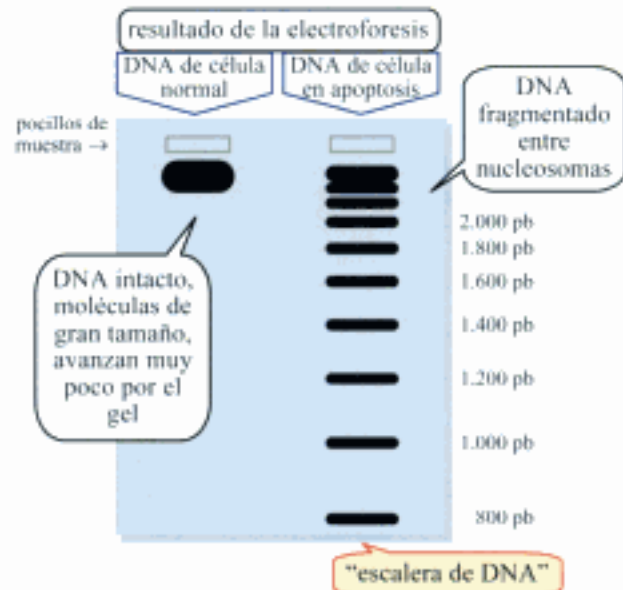
La apoptosis se controla mediante proteínas implicadas en varios niveles de señalización celular, codificadas por protooncogenes y genes oncosupresores. Al igual que el ciclo celular, el control de la apoptosis puede alterarse, generando situaciones patológicas entre las que se incluye el cáncer. Se conocen ya algunos genes que codifican proteínas reguladoras de la apoptosis y su alteración por mutaciones.

El **control positivo** consiste en el inicio de la cascada de caspasas; una de las principales señales inductoras es la liberación al citoplasma de los componentes de las mitocondrias rotas, en especial el **citocromo c** (integrante de la cadena respiratoria). Éste se une a la proteína **Apaf** (factor activador de las proteasas apoptóticas) formando un complejo que activa la procaspasa iniciadora. Algunas proteínas, como **Bax**, estimulan estas etapas de inicio; en ciertos cánceres Bax está inactiva, anulando su efecto de control positivo. Este control positivo de la apoptosis es análogo al ejercido en la proliferación celular por la inducción en cascada de la actividad proteína quinasa de los complejos Cdk-ciclina (pág. 426).

El **control negativo** de la apoptosis, por freno de la activación de las caspasas, viene determinado por proteínas como **Bcl-2** y **Bcl-x**, que aseguran que el proceso apoptótico permanezca bloqueado bajo condiciones normales. En algunos cánceres se sobreexpresa Bcl-2, reduciendo aún más la apoptosis; en otros, se sobreexpresan proteínas que bloquean las caspasas, como las denominadas **survivinas**. En general, este control es similar al ejercido sobre el ciclo celular por la activación de proteínas que inhiben la actividad proteína quinasa de los complejos Cdk-ciclina.

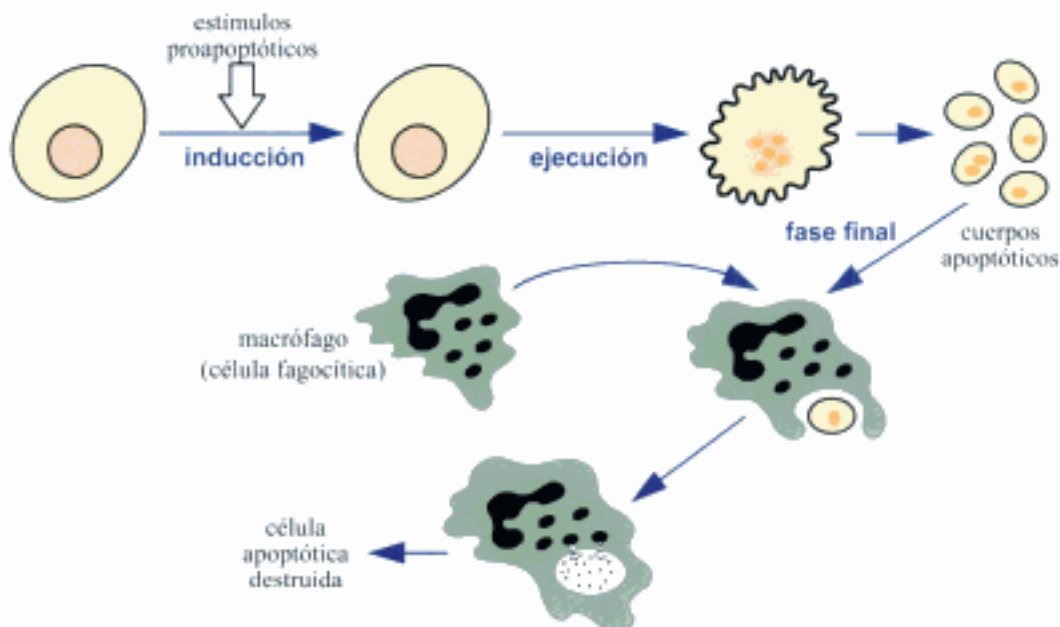
b) Nucleasas

Se encargan de la rotura del DNA cromosómico, generalmente en sus posiciones más accesibles, el DNA espaciador o internucleosómico, de forma que se originan fragmentos cuyo tamaño es múltiplo de 200 pb (longitud aproximada del DNA de cada nucleosoma, pág. 86). Por ello, uno de los síntomas de apoptosis es la observación, al estudiar el DNA por electroforesis, de un patrón de bandas en "escalera" (*DNA ladder*).



30.7.1.3 Fase final de la apoptosis

Las protuberancias externas de la célula, formadas durante la etapa de ejecución, terminan por separarse formando los llamados **cuerpos apoptóticos**; éstos son característicos y constituyen una de las diferencias respecto a la necrosis celular. Los cuerpos apoptóticos son reconocidos y fagocitados por otras células, tanto vecinales o del propio tejido como especializadas (macrófagos), de forma que toda la célula apoptótica se destruye sin que se llegue a liberar material celular al medio y, en consecuencia, sin la aparición de una reacción inflamatoria (a diferencia de la necrosis, donde la célula revienta, liberando su contenido y provocando esa respuesta). Obviamente, todo este proceso va acompañado de otros cambios bioquímicos.



30.7.2 Relación entre ciclo celular y apoptosis: p53 como punto de enlace

Como se ha indicado previamente, el gen oncosupresor p53 codifica la proteína nuclear p53, reguladora de la transcripción de ciertos genes relacionados con la progresión del ciclo celular (pág. 431). Por otra parte, p53 provoca la apoptosis, cuando las lesiones del DNA son imposibles de reparar, bien sea por su extensión o por defectos en los sistemas de reparación. Se elimina así la posibilidad de que se transmita el DNA gravemente alterado (con gran número de mutaciones, con reorganizaciones cromosómicas o con aneuploidia, por ejemplo). Los mecanismos por los cuales p53 produce esa activación de la apoptosis son múltiples y se conocen aún de forma incompleta, pero incluyen tanto la activación de la transcripción de genes como actividades no transcripcionales de la proteína p53. La alteración de estas funciones de p53, debido a la mutación de su gen, impide la entrada de la célula en apoptosis, lo que, unido a la desaparición de la función de freno del ciclo celular, tiene como consecuencia que las células continúen proliferando aunque posean daños en el DNA o defectos genéticos, aumentando así la probabilidad de que éstos se acumulen y favoreciendo la aparición de un cáncer.

30.8 HERENCIA DEL CÁNCER

Las alteraciones inducidas en el cáncer se transmiten en la mitosis de unas células a otras (células somáticas), pero no se transmiten por herencia de padres a hijos. Para ello habría de existir una predisposición o susceptibilidad a la mutación en la línea germinal, una **susceptibilidad hereditaria**. Existen, sin embargo, casos de cáncer hereditario, debidos a la mutación de protooncogenes o genes oncosupresores en el óvulo o el espermatozoide de los padres. La causa y mecanismo molecular del desarrollo del cáncer hereditario son las mismas que para el cáncer somático, sólo difieren en el tipo de célula donde se inicia (germinal o somática, respectivamente). Por otra parte, una vez heredado el cáncer, el individuo afectado se convierte en portador y lo transmitirá a sus descendientes de acuerdo con las reglas de herencia mendeliana y dependiendo del carácter dominante o recesivo del gen alterado causante del cáncer (por lo común, dominante si es un oncogén y recesivo si se trata de un gen oncosupresor mutado, págs. 417 y 419).

Uno de los cánceres hereditarios mejor conocidos y con especial trascendencia es el de mama, que tiene este carácter en un 5-10% de los casos. Su causa principal es la mutación de los genes oncosupresores BRCA1 del cromosoma 17 o BRCA2 del 13 (de **breast cancer**, cáncer de mama). El primero se expresa en mama y ovario, codificando un factor de transcripción con un motivo dedo de zinc (pág. 63) en su región N-terminal. Al tratarse de un gen oncosupresor, la transmisión hereditaria se debe generalmente a un alelo mutado, que sólo se manifestará en la hija desarrollando un tumor si también sufre mutación posteriormente el otro alelo, heredado como forma normal (pág. 419). La presencia de BRCA1 anormal se asocia en mujeres a un 50-80% de probabilidad de padecer cáncer de mama y a una predisposición al cáncer de ovario.

Existen otros defectos genéticos hereditarios, donde se transmite la predisposición a desarrollar un tipo específico de cáncer: xeroderma pigmentosa (pág. 363), enfermedad de Li-Fraumeni, cáncer cerebral, sarcoma, leucemia, ataxia telangiectasia, linfoma no hodgkiniano, leucemia linfocítica aguda, etc. Generalmente, la aparición del cáncer se demora bastante, pues es necesario que se acumulen mutaciones somáticas adicionales.

30.9 MARCADORES TUMORALES

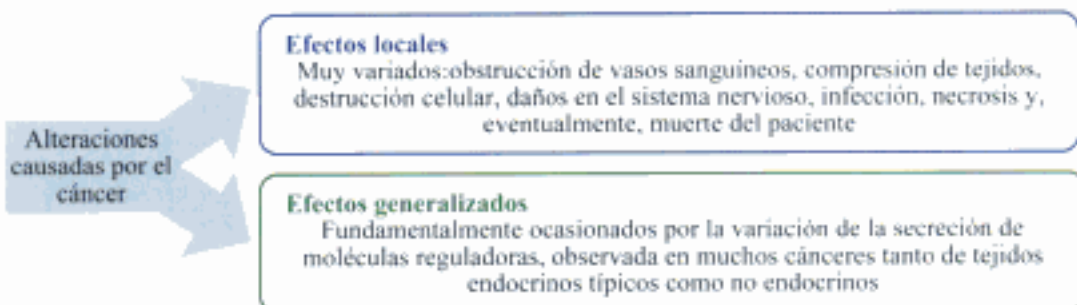
Las alteraciones de las células tumorales pueden considerarse a distinta escala. Lo tradicional y más utilizado en clínica son aquéllas que, aun sin estar directamente implicadas o ser responsables del proceso neoplásico, son consecuencia de la existencia del tumor o incluso de su evolución o respuesta a una determinada terapia. Con frecuencia, afectan a componentes específicos de la propia célula o de células vecinas. Estas alteraciones constituyen la base de los llamados **marcadores tumorales y tisulares**, establecidos con fines clínicos y relacionados, en general, con el grado de malignidad del tumor. Se observan tanto *in vitro* (por ejemplo en células en cultivo infectadas con virus oncogénicos en el proceso llamado *transformación maligna*) como *in vivo*, en numerosos casos clínicos.

El uso de los marcadores tumorales sólo es válido si para su determinación se cumplen ciertos requisitos analíticos. En cuanto a sus posibilidades de utilización clínica, deben hacerse algunas consideraciones básicas. Por un lado, raramente se emplean los datos de marcadores para establecer un diagnóstico, sino que éste ha de basarse en datos clínicos, radiológicos, biopsias, etc.; en general, los marcadores sólo sirven para confirmar la evidencia del cáncer. Tampoco deben utilizarse marcadores para buscar cáncer en una población (salvo excepciones, como en poblaciones de alto riesgo). En cuanto al pronóstico de una enfermedad, se exige una correlación entre la concentración del marcador y la masa del tumor. Son adecuados para el seguimiento del cáncer, en cuyo caso debe confirmarse la eficacia del tratamiento siguiendo la concentración del marcador a lo largo del tiempo, para averiguar, por ejemplo, si el aumento de un marcador es indicio de recurrencia de la enfermedad y se requiere una terapia más enérgica o un cambio de estrategia.

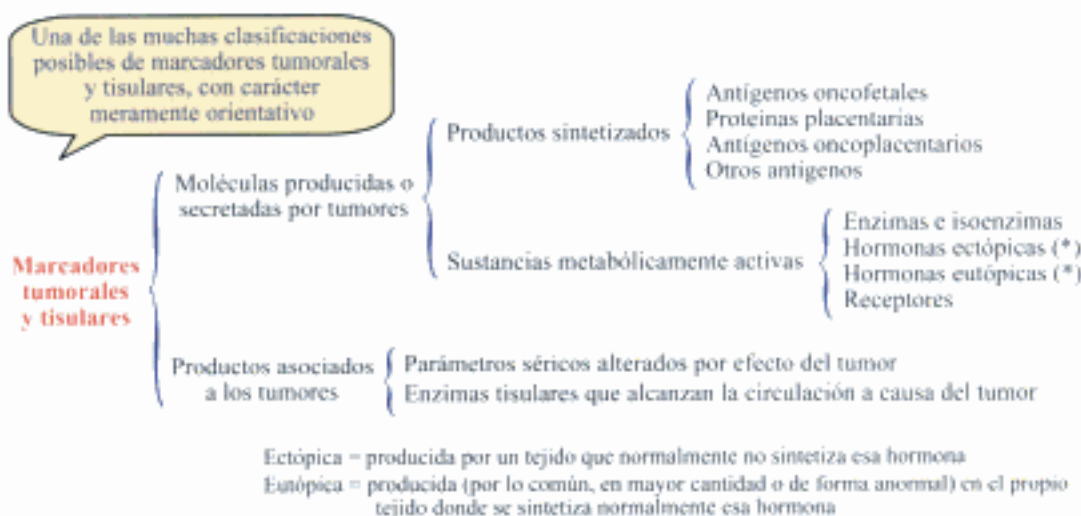
Las alteraciones de las células cancerosas son tan variadas que pueden estudiarse desde distintos puntos de vista: locales y generalizadas, tumorales y tisulares y, finalmente, genotípicas o moleculares.

30.9.1 Alteraciones locales y generalizadas

El proceso canceroso ejerce una influencia directa sobre la propia célula y su vecindad (efecto local) o sobre la totalidad del organismo (efecto generalizado o sistémico).



30.9.2 Alteraciones tumorales y tisulares



30.9.2.1 Marcadores tumorales

Se da este nombre a cualquier sustancia relacionada con la existencia o ausencia de un tumor en líquidos biológicos, en especial el plasma o el suero. Su naturaleza molecular es muy heterogénea: las más comunes tienen función hormonal o enzimática, o presentan características antigénicas. Su aparición pueden deberse a alteraciones de la morfología celular, del citoesqueleto (microtúbulos, microfilamentos y filamentos intermedios), de la superficie de la membrana (antígenos asociados, pérdida de inhibición del crecimiento por contacto, receptores), de todo tipo de parámetros bioquímicos o de la dotación cromosómica (cariotipo anormal, tema 29). Son los marcadores que más se analizan desde un punto de vista general. Se pueden distinguir dos grupos de marcadores tumorales: los de secreción y las citocinas (o citoquinas).

a) Marcadores tumorales de secreción o marcadores "clásicos"

Pueden clasificarse en función de los factores (tisulares, metabólicos, fisiológicos, metodológicos...) que determinan su concentración en líquidos biológicos, principalmente la sangre. La lista de marcadores de secreción es enorme y, a efectos prácticos, también pueden clasificarse en función del material biológico en el que se pueden estudiar: marcadores séricos, urinarios, de líquido cefalorraquídeo y de metástasis óseas.

A continuación se describen algunos de los marcadores de secreción mejor caracterizados.

α -Fetoproteína (AFP) o antígeno oncofetal

Cáncer de hígado

Glicoproteína de cadena sencilla (591 aminoácidos), con estructura similar a la α_1 -globulina, codificada por un gen del cromosoma 4. Se detecta a alta concentración en suero y líquido amniótico del feto normal de 7 a 9 semanas, donde se sintetiza en saco vitelino e hígado, con una función similar a la de la albúmina en adultos: mantener la presión oncótica y transportar sustancias. También pasa al suero materno, a través de la placenta. Es específica para el diagnóstico precoz del cáncer hepático y de tumores germinales que afectan a componentes del saco vitelino.

Subunidad β de la gonadotropina coriónica (β -hCG)**Cánceres de la línea germinal**

Glicoproteína expresada en la placenta durante el embarazo, formada por una cadena α (92 aminoácidos), codificada en el cromosoma 6, y otra β (145 aminoácidos), expresada por 7 genes del cromosoma 19. Su misión es estimular el cuerpo lúteo, una vez implantado el óvulo fecundado. La subunidad β libre se detecta en suero y, en ocasiones, en líquido cefalorraquídeo. Sirve para diagnosticar embarazo normal, gestaciones patológicas y ectópicas, cánceres testiculares de origen germinal y coriocarcinomas; aumenta también en el síndrome de Down. Se mide por métodos muy variados, principalmente inmunoquímicos, tanto cualitativos como cuantitativos.

 β_2 -microglobulina (β_2 mc)**Linfomas y leucemias**

Polipéptido pequeño (99 aminoácidos) sintetizado en casi todos los tejidos, forma parte de un antígeno de superficie de todas las células nucleadas. Puede pasar en forma soluble a la circulación y se filtra en los túbulo renales. Se detecta en suero y en orina. Se eleva en síndromes proliferativos, como mielomas, linfomas y leucemias, y es en general de utilidad en todo proceso con disfunción renal tubular.

Antígeno carcinoembrionario (CEA)**Cáncer de colon**

Familia de glicoproteínas expresadas por genes del cromosoma 19q. Ejercen funciones en adhesión celular, reconocimiento de bacterias, iniciación temprana del tumor, capacidad de metástasis y regulación del efecto inmunológico. Se presenta en varios carcinomas y sus metástasis, así como en tejidos embrionarios (de ahí el nombre). Se determina en suero, como marcador heterogéneo, poco específico (asociado a muchos carcinomas); se usa más en clínica para tumores derivados del endodermo (colorrectales, estómago, pulmón y mama).

Antígeno carbohidrato 125 (CA 125)**Cáncer de mama**

Glicoproteína de elevado tamaño, reconocida por un anticuerpo monoclonal, presente en varios procesos no tumorales y tumorales (secretada en células epiteliales del ovario), y asociada inespecíficamente a tumores de mama, pulmón, páncreas y colorrectal. Es específica para el carcinoma no mucinoso ovárico, y un factor de pronóstico para el cáncer de pulmón no microcítico. Se utiliza como marcador de elección en tumores epiteliales de ovario (excepto los mucinosos).

Calcitonina**Carcinomas medulares de tiroides**

Hormona peptídica de tamaño pequeño (32 aminoácidos) segregada por las células C del tiroides (o parafoliculares). El aumento de su secreción produce un descenso de la concentración de calcio libre en plasma, debido a la inhibición de su resorción ósea y su reabsorción tubular. Es útil para el diagnóstico precoz y el seguimiento del cáncer medular de tiroides.

Fosfatasa ácida prostática (PAP)**Cáncer de próstata**

Enzima específica de órgano, concretamente la isoenzima 2, implicada en el metabolismo de fosfolípidos y presente en el epitelio acinoso prostático. Se detecta en líquido seminal. Se utiliza para detectar y controlar a pacientes con cáncer de próstata.

Antígeno asociado al carcinoma de células escamosas (SCC)**Cáncer de útero**

Glicoproteína aislada de carcinoma del cuello uterino, presente en suero y específica como marcador epidermoide.

Tiroglobulina (TGB)**Cáncer de tiroides**

Proteína rica en residuos de tirosina, muy abundante en el tiroides como precursora de síntesis y almacén de hormonas tiroideas. La sintetiza el tirocito o célula tiroidea, en varios grados de yodación. Se determina en sangre, estando aumentada en el bocio tóxico, la tiroiditis subaguda y la enfermedad de Basedow. Es específica para cánceres diferenciados como el cáncer papilar y del foliculo tiroideo.

Queratinas CYFRA 21.1**Carcinoma escamoso de pulmón**

Queratina del tipo 19, presente en el citoesqueleto celular. Su valoración en suero es sensible a hepatopatías crónicas, insuficiencias renales y carcinomas escamosos de pulmón. Es también un indicador inespecífico de proliferación celular.

Apéndices

Referencias históricas

Aportaciones a la biología molecular e ingeniería genética, especialmente en relación con la patología molecular

Esta relación no pretende ser una lista exhaustiva de todas las aportaciones científicas que, a título individual o colectivo, se han hecho hasta configurar la materia presentada en este libro, sino que simplemente se desea reflejar los orígenes de las principales aportaciones, con una mención específica para las personas e hitos más significativos en nuestro país. En su mayoría, los datos seleccionados han surgido a partir de la información consultada, original o de fuentes indirectas. Los datos se ordenan cronológicamente agrupados bajo siete apartados.

1. Naturaleza de la materia viva

- 1660s R. Hooke observa las celdillas microscópicas (*células*) del corcho
- 1739 Primera edición de la farmacopea española (*Farmacopea Matritense*)
- 1838 J. Berzelius sugiere el término *proteína* para ciertas sustancias nitrogenadas de células animales y vegetales
- 1840s **Comienza la teoría celular** de los seres vivos con T. Schwann y M.J. Schleiden
- R. Brown observa una región densa circular dentro de las células vegetales (*núcleo*)
- J.E. Purkinje observa por primera vez el citoplasma (*protoplasma*) en el interior de la célula
- 1850s R. Virchow demuestra que toda célula deriva de otra; surge el concepto de *división celular*
- 1866 G. Mendel describe los mecanismos de herencia de los caracteres (plasmados en sus famosas *Leyes*)
- F. Hoppe-Seyler ocupa la primera Cátedra de Bioquímica (*Química Fisiológica*), Univ. Tubinga (Alemania)
- 1869 J.F. Miescher descubre la *nucleína*, una sustancia ácida rica en fósforo (el DNA)
- 1876 O. y R. Hertwig: la *sustancia genética* de los progenitores se transmite en la fecundación
- 1881 E. Zacharias demuestra que los cromosomas contienen la *nucleína* descubierta por J.F. Miescher
- 1882 W. Flemming: fusión de gametos en la fecundación y posterior separación del cigoto en dos células
- 1886 Se introduce la Química Biológica en España, en el doctorado de la Facultad de Farmacia de la Universidad Central de Madrid (Laureano Calderón Arana)

2. Surge la Genética como ciencia

- 1889 R. Altmann purifica un DNA libre de proteínas y acuña para la *nucleína* el término *ácido nucleico*
- 1896 Alfred Nobel dona su fortuna para crear los premios que llevan su nombre (se empiezan a conceder en 1901)
- 1900 C.E. Correns, H. de Vries y E.T.E. Seysseneg redescubren las leyes de Mendel, al cabo de 35 años
- 1902 A. Garrod analiza la *alcaptonuria* como alteración del metabolismo; surge el concepto de diversidad genética o polimorfismo
- W. Sutton concluye que los genes son unidades físicas bien definidas y localizadas en los cromosomas
- 1903 Primer libro de Bioquímica en España: *Tratado de Química Biológica*, de José Rodríguez Carracido
- 1906 W. Bateson observa el significado genético de la consanguinidad y acuña el término *Genética*
- 1906 F. Galton y A. Garrod inician la genética médica, con el concepto de errores congénitos del metabolismo
- Santiago Ramón y Cajal, primer premio Nobel español (con C. Golgi), por su trabajo sobre la estructura del sistema nervioso
- Se funda en España la *Junta para la Ampliación de Estudios*, bajo la presidencia de S. Ramón y Cajal
- 1909 T.H. Morgan estudia las mutaciones genéticas, los rasgos ligados al sexo y la función de los cromosomas
- W. Astbury utiliza por primera vez el término *Biología Molecular*, aunque sin su significado actual
- P. Johansen introduce el término *gen* para las unidades de la herencia con carácter transmisible
- 1910 A. Kossel, premio Nobel por sus estudios sobre la química de las células (proteínas y ácidos nucleicos)
- 1919 Ekery acuña el término *Biotecnología* para describir la "interacción de la biología y la tecnología"

- 1926 Premio Nobel para T. Svedberg por emplear la ultracentrifugación analítica para determinar masas moleculares
A. Muller (premio Nobel en 1946) demuestra en moscas que los rayos X causan mutaciones
- 1928 F. Griffith demuestra el proceso de *transformación bacteriana* (transferencia de ácido nucleico)
- 1929 A. Levine determina que la 2-desoxirribosa es el azúcar presente en los nucleótidos del DNA
- 1930 K. Landsteiner, premio Nobel por descubrir y tipificar los grupos sanguíneos humanos
- 1931 O.H. Warburg, premio Nobel por el estudio del proceso de respiración celular
- 1933 T.H. Morgan, premio Nobel por descubrir la función de los cromosomas como portadores de la herencia
- 1934 A. Felling describe la fenilcetonuria, una disenzimía congénita
J.D. Bernal y D. Crowfoot: primer estudio de una proteína mediante difracción de rayos X (pepsina cristalizada)
- 1937 J. Waldeström describe la macroglobulinemia y la primera variedad de porfiria, la porfiria cutánea tarda
- 1938 W. Astbury y Bell encuentran un espaciado regular de 0,34 nm a lo largo de las fibras del DNA
A. Huxley: la biotecnología llegará a ser más importante que la ingeniería mecánica y la química
- 1939 Se crea en España el *Consejo Superior de Investigaciones Científicas* (José María Albareda y Manuel Lora-Tamayo)

3. La bioquímica alcanza su madurez

- 1941 B. McClintock (premio Nobel en 1983) muestra que los transposones cambian de lugar en el genoma
G.W. Beadle y E.L. Tatum establecen la hipótesis un gen – una enzima
- 1943 G. de Hevesy, premio Nobel por la aplicación de los isótopos radiactivos a las reacciones metabólicas
H. Hinselmann publica una monografía sobre el diagnóstico de cáncer de útero mediante frotis vaginal
- 1944 O.T. Avery, C. MacLeod y M. McCarty: el DNA es el material portador de la información genética
Se introduce la Bioquímica en los estudios de licenciatura; comienza la promoción de bioquímicos españoles (Ángel Santos Ruiz)
- 1946 S. Takahara identifica la primera disenzimía silenciosa, la acatalasemia o carencia de catalasa
- 1947 Se crea en Argentina el *Instituto de Investigaciones Bioquímicas-Fundación Campomar*. Bernardo A. Houssay (premio Nobel en 1947) y Luis F. Leloir (director del *Instituto* durante 40 años y premio Nobel en 1970) constituyen un germen de la investigación bioquímica en Hispanoamérica.
- 1949 M.L. Barr y E.G. Bertram descubren la *cromatina sexual* en células de animales hembra
- 1950 D. Bevis describe la *amniocentesis* para detectar en el feto la incompatibilidad de factor Rh
- 1950s Se conoce ya gran parte de la composición química del DNA (azúcar, enlaces fosfodiéster y bases)
R. Franklin y M.H.F. Wilkins: primeros datos precisos por rayos X sobre la estructura del DNA
J.C. Kendrew y M.F. Perutz dilucidan mediante difracción de rayos X la estructura de mioglobina y hemoglobina (premio Nobel en 1962)
- 1951 L. Pauling descubre la base molecular de la anemia falciforme; surge el concepto de *enfermedad molecular*
H.G. Hers estudia las enfermedades lisosomales, disenzimías asociadas a orgánulos celulares
E. Chargaff establece la relación equimolecular existente entre purinas y pirimidinas en el DNA
- 1952 M. Chase y A.D. Hershey confirman en virus que el DNA es el portador de la información genética
Se crean la *Asociación Española de Biopatología Clínica* (AEBC) y la *Sociedad Española de Ciencias Fisiológicas* (SECF)

4. Se inicia la Biología Molecular con el estudio de la estructura del DNA

- 1953 F.H.C. Crick y J.D. Watson proponen la estructura en doble hélice para el DNA (premio Nobel en 1962)
- 1956 A. Levan y J.H. Tjio fijan en 46 el número de cromosomas humanos
Se descubren la RNA polimerasa y el mRNA; el mensaje genético requiere secuencias de pares de bases
F. Jacob y J. Monod demuestran la existencia de genes estructurales y genes reguladores
- 1957 Inicia sus actividades científicas el *Centro de Investigaciones Biológicas* (CIB); proyección de la Bioquímica hacia la Medicina (Alberto Sols y Carlos Asensio); se propone el término *Medicina Molecular*
- 1958 G.W. Beadle, J. Lederberg y E.L. Tatum, premio Nobel por estudios de recombinación génica y organización del material genético de bacterias
M. Meselson y F.W. Stahl demuestran por ultracentrifugación el carácter semiconservador de la replicación
A. Kornberg establece el mecanismo de la replicación catalizado por la DNA polimerasa
M. Gautier, J. Lejeune y R. Turpin: el mongolismo está causado por un cromosoma supernumerario
- 1959 A. Kornberg y S. Ochoa, premio Nobel por sintetizar *in vitro* polinucleótidos de RNA y DNA

- 1950-1960 Expansión de la Bioquímica en las Facultades de Farmacia y Ciencias de Barcelona (Fernando Calvet y Jorge Bozal; Vicente Villar Palasí y Manuel Rosell), Santiago de Compostela (José Antonio Cabezas), Granada (Federico Mayor), Sevilla (Manuel Losada) y Madrid (Ángel Martín Muñoz)
- 1960 Comienza el estudio del mecanismo de acción de los genes
P. Berg clona por primera vez el DNA
Se introduce la hibridación de los ácidos nucleicos
- 1961 F. Jacob y J. Monod (premio Nobel en 1965): la información genética fluye del DNA a través de un mRNA
H. Dintzis demuestra que la traducción discurre en sentidos 5'→3' del mRNA y amino→carboxilo de la proteína
M. Lyon muestra el proceso de *inactivación X* en el embrión de mamíferos hembra
M. Nirenberg muestra las primeras correspondencias entre codones y aminoácidos, sintetizando proteínas a partir de un DNA artificial
S. Brenner y F.H.C. Crick: la secuencia de codones fija la de aminoácidos; el código no se solapa ni se puntúa
M. Meselson y J. Weigle: la recombinación genética se produce por ruptura y nueva unión del DNA
- 1962 F.H.C. Crick, J.D. Watson y M.H.F. Wilkins, premio Nobel por sus estudios de estructura molecular del DNA
S. Benzer: el aminoacil-tRNA se incorpora en la síntesis proteica según la especificidad de su anticodón
- 1963 J. Cairns demuestra la síntesis simultánea de ambas hebras durante la replicación del DNA
Comienza a utilizarse el *ensayo de Guthrie* para el diagnóstico de la fenilcetonuria
Se funda la *Sociedad Española de Bioquímica* (actualmente *de Bioquímica y Biología Molecular*)
- 1964 R. Holliday propone un modelo para la recombinación homóloga
M. Roger consigue por primera vez implantar un gen en una célula anfitriona
- 1965 R. Holley determina la primera secuencia polinucleotídica natural (tRNA^{phe})
S. Benzer elabora el primer mapa detallado de un gen
W. Arber: base bioquímica de los procesos de restricción y modificación inducidos por el anfitrión
F. Jacob, A. Lwoff y J. Monod, premio Nobel por sus estudios del control genético de enzimas y de virus
- 1966 F.H.C. Crick propone el *balanceo* de la base 5' del anticodón durante la traducción
Se observa el paso de materiales a través del poro nuclear (por ejemplo, ribosomas)
- 1967 Se inicia en España la Biología Molecular (Antonio García-Bellido, David Vázquez y Eladio Viñuela), germen del *Centro de Biología Molecular* (CBM, inaugurado en 1975 y actualmente *C.B.M. "Severo Ochoa"*)
- 1968 N. Nomura y P. Traub consiguen ensamblar ribosomas *in vitro* a partir de sus componentes
R.W. Holley, H.G. Khorana y M.W. Nirenberg, premio Nobel por la interpretación del código genético
- 1969 Gran impulso desde el extranjero de la Bioquímica y Biología Molecular en España, por parte de Severo Ochoa, Santiago Grisolia y Francisco J. Ayala, entre otros
M. Delbrück, A.D. Hershey y S.E. Luria, premio Nobel por el mecanismo de replicación y la estructura genética de virus
Se inician en España los programas de Patología Molecular con la creación (Federico Mayor) del *Centro de Investigación de Alteraciones Moleculares y Cromosómicas* (CIAMYC)

5. Comienza la Ingeniería Genética, como tecnología del DNA recombinante

- 1970 R. Johnson y P. Rao sugieren los dos puntos de control (G₁/S y G₂/M) en el ciclo de división celular
Se observan en micrografías electrónicas los procesos de transcripción y traducción
R.J. Britten y D.E. Kohne diseñan las técnicas de reasociación del DNA eucariótico repetitivo
H.O. Smith y K.W. Wilcox descubren la primera enzima de restricción
D. Baltimore y M. Temin: los virus RNA pasan información genética a su anfitrión
G. Blobel (premio Nobel en 1999) y D. Sabatini proponen la *secuencia señal* para el tráfico de proteínas
- 1971 E.W. Sutherland, premio Nobel por sus estudios del mecanismo de acción de hormonas
Inicia sus actividades el *Instituto de Biología Fundamental* de Barcelona (Juan Oró, Jaime Palau y José Antonio Subirana)
- 1972 Se crea la *Sociedad Española de Genética*
- 1973 1ª *Reunión de Asilomar*: surge la preocupación ética por la manipulación mediante DNA recombinante
- 1974 R. Kornberg: el *nucleosoma* es la subunidad básica de empaquetamiento del DNA
A. Claude, C. de Duve y G.E. Palade: premio Nobel por la estructura y organización funcional de la célula
N.K. Jerné, G.J.F. Köhler y C. Milstein (premio Nobel en 1984) describen la *técnica del hibridoma* para producir anticuerpos monoclonales

- 1975 Estructura de tRNAs: anticodón (tripleto en el asa media) es complementario al codón (tripleto en el mRNA)
E.M. Southern desarrolla su técnica de hibridación de ácidos nucleicos. Con ella, A.M. Dozy, M.D. Golbus, Y.W. Kan y S. Mitchell hacen el primer diagnóstico prenatal de una enfermedad genética (la α -talasemia)

6. Se inicia el diagnóstico molecular

- 1975 D. Baltimore, R. Dulbecco y H.M. Temin, premio Nobel por mecanismos de producción de tumores por virus
Se crean la *Sociedad Española de Química Clínica* (SEQC; desde 1994 es la *Sociedad Española de Bioquímica Clínica y Patología Molecular*) y la *Asociación Española de Farmacéuticos Analistas* (AEFA)
- 1977 W. Gilbert y A.M. Maxam describen el primer método de secuenciación del DNA
R. Roberts y P. Sharp observan que el RNA se transcribe a partir de fragmentos discontinuos de DNA
R. Flavell y A. Jeffreys describen los intrones (secuencias intercaladas) en genes de eucariotas
La empresa Genentech patenta el proceso de clonación del DNA desarrollado por H.W. Boyer y S. Cohen, produciendo hormona del crecimiento en bacterias modificadas por ingeniería genética
Se crean en España los títulos oficiales de *Especialista en Análisis Clínicos* y *Bioquímica Clínica*
- 1978 W. Arber, D. Nathans y H.O. Smith, premio Nobel por descubrir las enzimas de restricción y aplicarlas a problemas de genética molecular
Se produce insulina humana por clonación de su gen en bacterias
A. Rich observa el Z-DNA por difracción de rayos X de cristales de hexanucleótidos ricos en GC
S.M. Tilghman visualiza los intrones en el gen de la globina
G. Cooper, R. Weinberg y M. Wigler descubren los oncogenes
- 1979 R. White y A. Wyman identifican el polimorfismo en la longitud de fragmentos de restricción (PLFR)
D. Lane y A. Levine descubren la proteína nuclear p53, sintetizada por un gen supresor de tumores
- 1980 W. Gilbert y F. Sanger, premio Nobel por sus métodos de secuenciación de DNA
P. Berg, premio Nobel por sus estudios bioquímicos de ácidos nucleicos (en especial, DNA recombinante)
C. Weismann logra producir interferón humano en bacterias
Se transfiere DNA a animales modelo mediante retrovirus
Primer diagnóstico de una enfermedad hereditaria, la anemia falciforme
- 1982 Genentech-Lilly comercializa la insulina humana (*humulina*), primer fármaco recombinante
R. Laskey: primera señal de localización celular que permite la entrada de proteínas al núcleo
A. Klug, premio Nobel por la aplicación de la microscopía electrónica a los complejos proteína-ácido nucleico
M. Barbacid, V. Notario y E. Santos identifican el primer oncogén humano, inductor de muchos tumores
- 1983 K. Mullis: técnica de la reacción en cadena de la DNA polimerasa (PCR) para producir DNA en cantidad
F. Gusella: marcador genético para la *enfermedad de Huntington* y diseño de un ensayo genético
B. McClintock, premio Nobel por el descubrimiento de los *transposones*, que se trasladan de un sitio a otro en el genoma
- 1984 R. Higuchi y A. Wilson: primera replicación de genes de una especie extinguida

7. Comienza la Genómica

- 1985 El *Departamento de Energía* de EE.UU. (DOE) plantea métodos para secuenciar genes de interés clínico
Comercialización de proteínas recombinantes: hormona de crecimiento (*Protropin*), interferón (*Roferón*), etc.
A. Jeffreys desarrolla la *huella genética* para identificar personas a partir de su material biológico
Se automatiza la secuenciación del DNA
Se inician las actividades de la *Sociedad Española de Biotecnología* (constituida formalmente en 1985)
- 1986 Se autoriza una vacuna para la hepatitis B producida por ingeniería genética
Se crea en los *Institutos Nacionales de la Salud* (NIH) de EE.UU. el *Comité Consultivo del DNA Recombinante*
H. Roderick acuña el término *genómica*: ciencia y métodos para localizar genes y hallar su secuencia
V.A. McKusick recopila unas 4.000 alteraciones genéticas hereditarias distintas
- 1987 Se descubre que la metilación del DNA en eucariotas silencia su transcripción
Se propone el inicio del *Proyecto Genoma Humano* (PGH), coordinado por la *Organización del Genoma Humano* (HUGO), con sedes en Londres, Bethesda y Tokio. En EE.UU. se crea la *Oficina de Investigación del Genoma* (luego, *Centro Nacional de Investigación del Genoma Humano*), presidida por J.D. Watson hasta 1992.
Se desarrollan los cromosomas artificiales de levadura (vectores YAC)

- 1988 J.W. Black, G.B. Elion y G.H. Hitchings, premio Nobel por el tratamiento con fármacos de varias enfermedades
- 1989 S. Altman y T.R. Cech, premio Nobel por demostrar propiedades catalíticas en algunos RNAs
J.M. Bishop y H.E. Varmus, premio Nobel por el descubrimiento de los oncogenes retrovirales
- 1990 C. Venter desarrolla una estrategia para localizar las EST (etiquetas de secuencias expresadas)
- 1991 M.C. King asocia un gen del cromosoma 17 con la aparición de cáncer de mama familiar
- 1992 E.H. Fischer y E.G. Krebs, premio Nobel por los mecanismos de fosforilación reversible de las proteínas
Se obtienen mapas físicos del cromosoma 21 (D. Cohen) y del cromosoma Y (D. Page)
H. Noller demuestra que la actividad peptidiltransferasa se localiza en un rRNA del ribosoma
- 1993 P.A. Sharp y R.J. Roberts, premio Nobel por observar genes discontinuos (intrones+exones) en eucariotas
W. Akey elabora el primer modelo tridimensional del *complejo molecular del poro nuclear*
A. Rosses establece la conexión entre el gen *ApoE* y la enfermedad de Alzheimer
K.B. Mullis, premio Nobel por el método de reacción en cadena de la DNA polimerasa (PCR)
M. Smith, premio Nobel por la mutagénesis dirigida mediante oligonucleótidos
- 1994 A.G. Gilman y M. Rodbell, premio Nobel por descubrir las proteínas G y su papel en la transducción de señales
- 1996 Se secuencian los genomas de la bacteria más pequeña, *Mycoplasma genitalium*, y de la levadura de cerveza, *Saccharomyces cerevisiae*
Affimetrix comercializa un biochip para analizar las mutaciones del virus del SIDA
- 1997 Se hace posible la determinación de cromosomopatías en células sanguíneas fetales
Conclusión del mapa físico de alta resolución de los cromosomas X y Y
- 1998 Secuenciación completa del genoma del gusano *Caenorhabditis elegans*
El gen de bleomicina-hidroxilasa está ligado a la enfermedad de Alzheimer, independientemente del gen *ApoE*
C. Venter crea Celera con fondos privados para secuenciar el genoma humano en 3 años, y descifra el genoma de microorganismos patógenos y de interés comercial y farmacéutico
- 1999 R. Weinberg transforma *in vitro* una célula humana sana en cancerosa mediante mutaciones genéticas
Secuencia completa y mapa genético del complejo mayor de histocompatibilidad en el cromosoma 6
Un consorcio público (Europa, EE.UU., Japón) secuencian el 97% del primer cromosoma humano, el 22
Pruebas sanguíneas para el diagnóstico del cáncer colorrectal a través de guanilciclasa
- 2000 Bioquímicos, biólogos moleculares y genéticos ocupan puestos de responsabilidad en las *Reales Academias de Ciencias* (Ángel Martín Municio y Enrique Sánchez-Monge y Parellada) y de *Farmacología* (Julio Rodríguez Villanueva) y el *Instituto de España* (Margarita Salas Falgueras)
Mapa completo de la mosca *Drosophila melanogaster*
Los políticos de EE.UU. y Reino Unido piden el libre acceso a los datos del genoma
Se descifra el mapa genético de los cromosomas humanos 5, 16 y 19
Se secuencian los cromosomas 21q, que sorprende por los pocos genes que contiene
Se hace la presentación del genoma humano (95% secuenciado, conjuntamente por PGH y Celera)

Bibliografía

Se recoge una relación de libros de texto recientes, en castellano, adecuados para complementar el contenido de este libro.

- (1999) Diccionario Mosby de Medicina, Enfermería y Ciencias de la Salud, 5ª ed. Harcourt, Madrid
- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. y Watson, J.D. (1996) *Biología Molecular de la Célula*, 3ª ed. Omega, Barcelona
- Alberts, B., Bray, D., Johnson, S., Lewis, J., Raff, M., Roberts, K. y Walter, P. (1999) *Introducción a la Biología Celular*. Omega, Barcelona
- Anderson, S.C. y Cockayne, S. (1995) *Química Clínica*. Interamericana-McGraw-Hill, Madrid
- Baiget, M., Gallano, P. y Tizzano, E. (1995) *Técnicas de Biología Molecular*. Comité de Publicaciones de la Sociedad Española de Bioquímica Clínica y Patología Molecular, Barcelona
- Balcells, A. (1999) *La Clínica y el Laboratorio*, 18ª ed. Masson, Barcelona
- Cárdenas, J., Fernández, E., Muñoz, J. y Pineda, M. (1996) *Glosario de Biología Molecular*. Servicio de Publicaciones de la Universidad de Córdoba
- Chernicky, C.C. y Berger, B.J. (1999) *Pruebas Diagnósticas y de Laboratorio*. Interamericana-McGraw-Hill, Madrid
- Conn, E.E. y Stumpf, P.K. (1996) *Bioquímica fundamental*, 5ª ed. Limusa, México
- Cummings, M.R. (1996) *Herencia Humana: Principios y Conceptos*, 3ª ed. Interamericana-McGraw-Hill, Madrid
- Darnell, J., Lodish, H. y Baltimore, D. (1993) *Biología celular y molecular*, 2ª ed. Omega, Barcelona
- De Robertis, E.M.F., Hib, J. y Poncio, R. (1997) *Biología Celular y Molecular de E.P.D. de Robertis*. El Ateneo, Buenos Aires
- Devlin, T.M. (1999-2000) *Bioquímica: Libro de Texto con Aplicaciones Clínicas*, 3ª ed., 2 vols. Reverté, Barcelona
- Díaz Zagoya, J.C. y Hicks, J.J. (1995) *Bioquímica*, 2ª ed. Interamericana-McGraw-Hill, Madrid
- D'Ocón Navaza, M.C., García García-Saavedra, M.J. y Vicente García, J.C. (1999) *1000 Cuestiones sobre Técnicas de Laboratorio*. Paraninfo, Madrid
- Etienne-Decant, J. (1990) *Bioquímica Genética*. Masson, París
- Freifelder, D. (1991) *Técnicas de Bioquímica y Biología Molecular*. Reverté, Barcelona
- Fuentes Arderiu, X., Castiñeiras Lacambra, M.J. (1998) *Diccionario de Ciencias de Laboratorio Clínico*. McGraw-Hill Interamericana
- Fuentes Arderiu, X., Castiñeiras Lacambra, M.J., & Queraltó Compañó, J.M. (1998) *Bioquímica Clínica y Patología Molecular*, 2 vols. Reverté, Barcelona
- Garrido Pertierra, A. (1991) *Fundamentos de Química Biológica*. Interamericana-McGraw-Hill, Madrid
- Gaw, A., Cowan, R.A., O'Reilly, D., Stewart, M.J. y Shepherd, J. (2000) *Bioquímica Clínica Ilustrada*, 2ª ed. Harcourt, Madrid
- González de Builrago, J.M., Arilla Ferreiro, E., Rodríguez-Segade, M. y Sánchez Pozo, A. (1998) *Bioquímica Clínica*. McGraw-Hill Interamericana, Madrid
- González Duarte, R. (1991) *Biología de Aplicación Farmacéutica*. Doyma, Barcelona
- Griffiths, A.J.F., Gelbart, W.M., Miller, J.H. y Lewontin, R.C. (2000) *Genética Moderna*. Interamericana-McGraw-Hill, Madrid
- Hernández Bronchud, M. (1992) *La Industria del DNA. Principios Básicos y Potencial Futuro*. Editorial MCR, Barcelona
- Herrera, E. (1991) *Bioquímica. Vol I: Aspectos Estructurales y Metabólicos. Vol II: Bioquímica Molecular y Bioquímica Fisiológica*, 2ª ed. Interamericana-McGraw-Hill, Madrid
- Herrera, E. (1993) *Elementos de Bioquímica*. Interamericana-McGraw-Hill, Madrid
- Hicks, J.J. (1999) *Bioquímica*. Interamericana-McGraw-Hill, Madrid
- Horton, H.R., Moran, L.A., Ochs, R.S., Rawn, J.D. y Scrimgeour, K.G. (1995) *Bioquímica*. Prentice Hall
- Izquierdo Rojo, M. (1995) *Biología Molecular del Cáncer*. Síntesis, Madrid
- Izquierdo Rojo, M. (1999) *Ingeniería Genética y Transferencia Génica*. Pirámide, Madrid

- Jones, E. y Morris, A. (1999) Curso "Crash de Mosby": Lo Esencial en Célula y Genética. Harcourt, Madrid
- Jorde, L.B., Carey, J.C., Bamshad, M.J. y White, R.L. (2000) Genética Médica, 2ª ed. Harcourt, Madrid
- Junqueira, I.C. y Carneiro, J. (1998) Biología Celular y Molecular, 6ª ed. McGraw-Hill Interamericana, Madrid
- Karp, G. (1998) Biología Celular y Molecular. McGraw-Hill Interamericana, México
- Lacadena, J.R. (1999) Genética General. Síntesis, Madrid
- Lehninger, A.L., Nelson, D.L. y Cox, M.M. (1993) Principios de Bioquímica, 2ª ed. Omega, Barcelona
- León Serrano, J. y García Lobo, J.M. (1990) Manual de Genética Molecular. Síntesis, Madrid
- Lewin, B. (2000) Genes VII. Oxford University Press, Oxford
- López Colomé, R. (2000) Bioquímica y Biología Molecular. McGraw-Hill Interamericana, Madrid
- Lozano Teruel, J.A., Galindo Cascales, J.D., García-Borrón Martínez, J.C., Martínez Liarte, J.M., Peñafiel García, R. y Solano Muñoz, F. (2000) Bioquímica y Biología Molecular para Ciencias de la Salud, 2ª ed. McGraw-Hill Interamericana, Madrid
- Lucas, M. (coord.) (1998) Diagnóstico Molecular en Genética Médica. Comité de publicaciones de la Sociedad Española de Bioquímica Clínica y Patología Molecular, Barcelona
- Macarulla, J.M. y Goñi, F.M. (1990) Bioquímica humana: curso básico, 2ª ed. Reverté, Barcelona
- Macarulla, J.M., Marino, A., y Macarulla, A. (1992) Bioquímica Cuantitativa Vol.II: Cuestiones sobre Metabolismo. Reverté, Barcelona
- Martínez Jarreta, M.B. (1999) La Prueba del ADN en Medicina Forense. Masson, Barcelona
- Mathews, C.K. y Van Holde, K.E. (1998) Bioquímica, 2ª ed. McGraw-Hill Interamericana, Madrid
- Medina, J.M., Sánchez de Medina, F. y Vargas, A. (1996) Bioquímica. Síntesis, Madrid
- Montgomery, R., Conway, T.W., Spector, A.A. y Chapel, D. (1998) Bioquímica: Casos y Texto, 6ª ed. Harcourt, Madrid
- Muñoz, A. (1997) Cáncer: Genes y Nuevas Terapias. Hélice, Madrid
- Murray, R.K., Granner, D.K., Mayes, P.A. y Rodwell, V.W. (2000) Bioquímica de Harper, 15ª ed. El Manual Moderno, México
- Navarro, F.A. (2000) Diccionario Crítico de Dudas Inglés-Español de Medicina. McGraw-Hill Interamericana, Madrid
- Oliva Virgili, R. (1996) Genoma Humano. Masson, Barcelona
- Paniagua, R. (1999) Biología Celular. McGraw-Hill Interamericana, Madrid
- Peña, A., Arroyo, A., Gómez, A. y Tapia, R. (1992) Bioquímica, 2ª ed. Limusa, México
- Pincheira, J. (2000) Daño, Reparación y Mutación del Material Genético. Publicaciones Técnicas Mediterráneo, Chile
- Puerta, M.J. (1999) Genética: Fundamentos y Perspectivas, 2ª ed. McGraw-Hill Interamericana, Madrid
- Quesada, A.R., Medina Torres, M.A. y Sánchez Jiménez, F. (coord.) (2000) Biología Molecular del Cáncer: Fundamentos y Perspectivas. Servicio de Publicaciones de la Universidad de Málaga
- Real Academia de Ciencias Exactas Físicas y Naturales (1999) Diccionario Esencial de las Ciencias. Espasa Calpe, Madrid
- Roitt, I., Brostoff, J. y Male, D. (2000) Inmunología, 5ª ed. Harcourt, Madrid
- Roskoski, R., Jr. (1998) Bioquímica. McGraw-Hill Interamericana, Madrid
- Ruiz Amil, M. (1992) Bioquímica Metabólica: Conceptos Fundamentales. Tebar Flores, Madrid
- Salve, M.L., Amich, S., Prieto, S. y Casas, A. (1994) Laboratorio Clínico de Bioquímica. Interamericana-McGraw-Hill, Madrid
- Singer, M. y Berg, P. (1995) Genes y Genomas: Una Perspectiva Cambiante. Omega, Barcelona
- Smith, C.A. y Wood, E.J. (1998) Biología Molecular y Biotecnología. Addison-Wesley Iberoamericana, Wilmington, U.S.A
- Smith, A.D., Datta, S.P., Smith, G.H., Campbell, P.N., Bentley, R., & McKenzie, H.A. (1997) The Oxford Dictionary of Biochemistry and Molecular Biology (ed. revisada 2000). Oxford University Press
- Strachan, T. y Read, A.P. (1999) Genética Molecular Humana. Omega, Barcelona
- Stryer, L. (1995) Bioquímica, 4ª ed. Reverté, Barcelona
- Thompson, M., McInnes, R.R. y Willard, H.F. (1996) Genética en Medicina, 4ª ed. Masson, Barcelona
- Tormo Garrido, A. (1998) Problemas de Genética Molecular. Síntesis, Madrid
- Voet, D. y Voet, J.G. (1992) Bioquímica. Omega, Barcelona
- Wallach, J. (1998) Interpretación Clínica de las Pruebas de Laboratorio, 3ª ed. Masson, Barcelona
- Weller, B. (1995) Diccionario Enciclopédico de Ciencias de la Salud. Interamericana-McGraw-Hill, Madrid

Índice analítico

Notas relativas a la ordenación alfabética de los términos:

Se han ignorado letras griegas, números y otros indicadores de posición del sustituyente (tales como *N*- y *O*-).

Por ejemplo, '2-**a**cetamidofluoreno' y '*N*-**a**cetilglucosamina' aparecen en la '**a**'.

Los ácidos se citan en la posición de su nombre, no en la '*a*'. Por ejemplo, '**f**usídico (ácido)'.

Las enfermedades específicas se citan por su nombre, no en la '*e*'. Por ejemplo, '**A**lzheimer (enfermedad de)'.

Los nombres de enzimas se citan completos, no como subentendidos de la primera palabra. Por ejemplo, 'DNA polimerasa' no forma parte de la entrada 'DNA'. En general se emplea el nombre común y sin guiones intermedios, siguiendo el uso internacional en inglés y la recomendación de 1994 de la Sociedad Española de Bioquímica Clínica y Patología Molecular.

A

- A1AT. Véase α_1 -antitripsina
- ABO (gen), 370. Véase también grupo sanguíneo ABO
- aaRS. Véase aminoacil-tRNA sintetasa
- absortividad (absorción óptica)
- de ácidos nucleicos, 166
 - de bases nitrogenadas, 166
 - de nucleósidos y nucleótidos, 28, 166
- variación en la desnaturalización del DNA, 165
- ABTS, 186
- acéntrico. Véase cromosoma
- 2-acetamidofluoreno, 355, 359
- N*-acetilgalactosamina, 330
- N*-acetilglucosamina, 330
- N*-acetilneuraminico (ácido), 330
- β -*N*-acetilhexosaminidasa. Véase hexosaminidasa A
- acetilación
- de aminoácidos, 326
 - de histonas, 264
 - y expresión génica, 264
- acetiltransferasa
- de histonas, 264
- N*⁶-acetiltransferasa, 326
- aciclovir, 13
- ácido úrico, 13
- ácidos, efecto sobre la desnaturalización, 164
- ácido nucleico, 1
- absorción en el ultravioleta, 36, 166
 - carácter hidrofílico, 34
 - carácter polianiónico, 34
 - componentes, 9
 - esqueleto, 32
- ácido nucleico
- estructura primaria, 32
 - extracción, 134, 135
 - función, 31
 - hidrólisis química, 35
 - localización, 31
 - niveles estructurales, 32
 - propiedades fisicoquímicas, 34
 - purificación, 134
 - reactividad, 34
 - representaciones lineales, 32
 - solubilidad, 134
 - tipos, 33
 - viscosidad, 34
- acilación de aminoácidos en proteínas, 334
- acortamiento de los extremos del DNA, 157, 158
- acridina, 259, 355. Véase naranja de acridina
- actina
- escisión por caspasas, 432
 - familia multigénica, 112
- actinomomicina D, 258
- activación
- de grupos funcionales, 178, 179
 - de aminoácidos, 302, 303, 307
- activador, 266
- actividad transcripcional, 91, 263
- ADA. Véase adenosina desaminasa
- adaptador, 195
- adenina, 12
- adenina fosforribosiltransferasa. Véase APRT
- S*-adenosilhomocisteína, 201, 273
- S*-adenosilmethionina, 22, 199, 200, 273, 278
- adenosina, 20
- como centro de ramificación, 281
- adenosina desaminasa
- deficiencia de, 396
- adenovirus, 431
- adhesión, 127
- alteración en apoptosis, 432
 - célula-matriz extracelular, 413-414
 - intercelular, 413-414
- A-DNA, 52
- geometría de los pares de bases, 54
 - modelos moleculares, 53, 54
- ADP-ribosilación, 336
- adriamicina, 259
- afinidad, 139
- métodos, 128
 - para extracción de mRNA, 280
 - por la cola de poli(A), 280
- aflatoxina, 355
- AFP, 435
- agar, 170
- agarosa, 138
- aglutinación, 370
- alcohol isoamílico
- en la extracción del DNA, 134
- aldolasa
- familia multigénica, 112
- alelos, 95, 96
- coheredados, 218
 - ligados, 220
 - segregados, 218
 - y polimorfismo, 365
- alfoide, 114
- alquilación, 355
- reparación, 360
- alquilguanina. Véase O-metilguanina
- alquiltransferasa, 357
- mecanismo de catálisis, 358
- alteraciones genómicas, 403
- Alu, 114, 324

- Alzheimer (enfermedad de)
relación con el plegamiento, 337
- α -amanitina, 251
estructura química, 258
- aminas aromáticas, 355
- aminoácidos
abreviaturas de 1 y 3 letras, 294
acetilación, 326
acilación, 334
activación, 302
asignación de codones, 290
básicos, 84
carboxilación, 327
fosforilación, 327
hidroxilación, 328
metilación, 328
modificación en las proteínas, 326
prenilación, 335
unión por enlace peptídico, 313
- aminoácido-tRNA ligasa. Véase
aminoacil-tRNA sintetasa
- aminoacilación del tRNA, 302
- aminoacil-AMP, 302, 303
verificación, 305
- aminoacil-tRNA, 289
2' y 3', 303
ubicación en el sitio A, 311
verificación, 305
- aminoacil-tRNA sintetasa, 302
características, 304
centros activos, 304
corrección de errores, 305
de clase I, 303, 304
de clase II, 303, 304
especificidad, 304
modelo molecular, 305
número necesario, 305
reconocimiento del aminoácido,
304, 305
- reconocimiento del tRNA, 304,
305
- aminoglicósido, 316, 317
- aminopterina en medio HAT, 223
- amnios, 120
- AMP cíclico, 29
- ampicilina, gen de resistencia,
209, 214
- amplificación genética.
Véase clonación
- amplificador, 266
- amp', 209, 214
- anafase, 99
- anclaje de proteínas a membranas
a través de lípidos, 334
- andamiaje nuclear.
Véase esqueleto nuclear
- anemia. Véase falciforme
y véase Falconi
- aneuploidia, 403, 404
mecanismo, 404, 405
- anfitrión. Véase célula
- angiogénesis, 414
- anomalía cromosómica.
Véase cromosomopatía
- anotación, 240
- antibiótico
de tipos I, II y III, 259
que inhibe la traducción, 316
resistencia, 206
Véase también α -amanitina,
ampicilina, puromicina,
rifampicina...
- anticodón, 74, 289
apareamiento con el codón, 296
número necesario, 297
relación con el número de
codones, 298
- anticuerpos
anti-digoxigenina, 186
aplicación en ensayo Western,
173
monoclonales, 127
que existen naturalmente, 369
- antifólico, 223
- antígeno
A, 372
asociado al carcinoma de
células escamosas, 436
B, 372
carbohidrato 125 (CA125), 436
carcinoembrionario, 436
celular de células en
proliferación. Véase PCNA
H, 372
oncofetal, 435
T del virus SV40, 431
- antioncogén. Véase oncosupresor
- antiparalelismo, 44
- antisentido, 78, 248
- α -antitripsina, 376
herencia, 396
isoforma Z, 376
mutaciones causantes, 376
- α -AT. Véase α -antitripsina
- antiviral, 21, 23
- Apaf (proteína), 433
- apareamiento de bases
codón-anticodón, 296
entre 3 bases, 73
incorrecto, 362
intracatenario, 70, 73
temporal, 248
Véase también
complementariedad
- apilamiento, 14, 48, 165
- apirimidínico, 35, 232, 354, 361
- apoptosis
concepto, 431
control negativo, 416, 433
control positivo, 418, 433
cuerpos apoptóticos, 433
ejecución, 432
equilibrio con proliferación, 414
etapas, 432, 433
fagocitosis, 433
fase final, 433
inducción, 432
iniciación por citocromo c, 433
papel de p53, 431
relación con proliferación, 431
- apoptosis
señales negativas, 432
señales positivas, 432
- APRT, 223
- apurínico, 35, 232, 354, 361
- AraA o arabinosiladenina, 21
- AraC o arabinosilicitosina, 21
- árbol genealógico, 385
nomenclatura y símbolos, 393
- Arf (gen y proteína), 421, 426
- ARS. Véase secuencia de
replicación autónoma
- ARS/CEN/TEL, 211
- arsénico, 355
- asa de desdoblamiento o
de desplazamiento.
Véase bucle D
- asbesto, 355
- ascorbato, 328
- aspartato aminotransferasa
como marcador tumoral, 437
- ataxia-telangiectasia, 363, 406
- ATP como cosustrato de la ligasa, 156
- ATPasa
dependiente de RNA, 255
- AttoPhos, 186
- autoiniciadora, 248
- automatización
de la PCR, 193
de la secuenciación enzimática
de DNA, 237
de la síntesis de
oligonucleótidos, 179
- autoradiografía, 170, 185, 186
- autosoma, 93
- Avery, MacLeod y McCarty, 2
- avidina, 129, 186
- ayuste, 280
alternativo, 276, 285
centros de, 280
mecanismo, 281
reconocimiento, 281
secuencias, 280
- ayustosoma, 280
- AZT. Véase zidovudina
- azúcares
activados con nucleótido, 332
componentes de ácidos
nucleicos, 11
configuración anomérica, 20
- azul de tripano, 129

B

- BAC, 209, 211
- bacteria
cápsula, 3
colonia, 170, 214
lisis *in situ*, 170
- bacteriófago, 210
 λ , 210
M13, 210
P1, 211
T2, 4
- baculovirus, 210

- balanceo, 296, 298, 304
 de guanina, 297
 de inosina, 297
 de uracilo, 296
- bandas cromosómicas
 G, Q, R y T, 91
- bando
 de alta resolución, 91
 de cromosomas, 90
 de profase, 91
 de prometáfase, 91
- barrera, métodos de, 124
- basal. Véase promotores
- bases, efecto sobre la
 desnaturalización, 164
- bases nitrogenadas
 absorción en el ultravioleta, 17, 166
 carácter ácido-base, 16
 complementariedad, 39
 componentes de ácidos
 nucleicos, 12
 derivados metilados, 13, 72,
 201, 272
 desaminación, 353
 eliminación selectiva, 232
 estructura, 12
 infrecuentes, 13
 metilación, 355
 modelos moleculares, 14
 numeración de los átomos, 12
 propiedades fisicoquímicas, 14
 tautomería, 14
 vías de síntesis, 223
- Bax (proteína), 433
- Bcl-1 (gen), 417
- Bcl-2 (gen y proteína), 417, 433
- Bcl-x (proteína), 433
- B-DNA, 39
- benzo[a]pireno, 355
- benzoiladenina, 179
- benzocitosina, 179
- BER. Véase escisión de bases
- berilio, 355
- biblioteca de DNA.
 Véase genotecas
- bidireccional. Véase replicación
- bifocal. Véase replicación
- bifuncionales, 355
- biochip, 387, 437
- bioinformática, 380
- biolística, 212
- biopsia, 119
- biotina, 196, 336
- bisfosfato, 11, 29
- blastocela, 119
- blastocisto, 119
- blastodermo, 119
- blastómero, 119
- Bloom (síndrome de), 406
- bloque de repetición.
 Véase DNA repetitivo
- bombardío, 212
- Bombay (fenotipo), 373
- brazos
 en la estructura del tRNA, 74
- BRCA (genes), 434
- BrdU o bromodesoxiuridina, 132
- bromouracilo, 355
- bromuro de etidio, 126, 138, 259
- bucle, 70
- bucle D, 257, 377
- burbuja, 70
 de transcripción, 253
 de transcripción, desplazamiento, 254
- bZIP (proteínas), 64
- C**
- C₂H₅. Véase dedo de zinc
- C₃. Véase dedo de zinc
- CA125 (antígeno), 436
- CAAT (secuencia o caja), 266, 269
- cabeza de martillo, 78
- cadmio, 355
- cafeína, 13
- CAK. Véase quinasa activadora
 de Cdk
- calcitonina, 436
- calmodulina, 328
- calpains, 342
- cAMP. Véase AMP cíclico
- cáncer
 angiogénesis, 414
 colorrectal hereditario sin
 poliposis, 363
 como enfermedad genética, 412
 concepto, 412
 de colon, 436
 de hígado, 435, 376
 de la línea germinal, 434, 436
 de mama, 434, 436
 de ovario, 434
 de próstata, 436
 de pulmón, 436
 de tiroides, 436
 de útero, 436
 etapas, 413
 germinal y somático, 434
 herencia, 434
in situ, 413, 414
 invasivo, 413
 origen monoclonal, 413
 origen por mutación, 413
 vascularización, 414
 y metilación del DNA, 273
 y telomerasa, 160
- CAP. Véase proteína receptora
 de AMP cíclico
- caperuza 5', 278
 papel protector, 278
 papel en el inicio de la
 traducción, 309, 310
 tipos 0, 1 y 2, 278
- cápsida vírica, 4
- cápsula bacteriana, 3
- carabina, 337-339
 actividad ATPasa, 339
 función, 340
 tipos, 339
- caracteres, 95
- β-carboxiaspartato, 327
- γ-carboxiglutamato, 327
- carboxilación
 de aminoácidos, 327
- carboxilasa, 327
- cariograma. Véase cariotipo
- cariotipo, 89
 análisis, 89
 de flujo, 225
 molecular, 172
 preparación, 89
- cascada
 de señales, 421
 de caspasas, 432
- caseína, fosforilación, 327
- casero (gen) 269
- caspasa, 432
 activación en cascada, 432
 ejecutora, 432
 inhibición, 433
 iniciadora, 432
- casquete. Véase caperuza
- catepsina, 341
- CBP, 309, 310, 319
- Cdc25, 427
- Cdk, 424
 activación, 426-428
 actividad quinasa, 425-427
 complejo con ciclina, 424, 426,
 427
 especificidad de sustrato, 424
 fosforilación, 427
 inhibidores, 426
 regulación del ciclo celular, 427
 síntesis, 425
 sustratos, 428
 tipos, 426
- Cdk1, 417
- Cdk2, inhibición por p21, 431
- cDNA, 196, 244
 como inserto, 205
 genotecas de, 216
 preparación, 206
 y EST, 228
- CEA, 436
- cebado al azar, 183
- cebador, 145
 eliminación, 156, 157
 empleo en PCR, 189
 Véase también primasa
- celadora. Véase carabina
- célula anfitriona, 203, 211
- célula germinal, 97, 99
 ciclo celular, 100
 primordial, 102, 103
- célula hospedadora. Véase célula
 anfitriona
- célula madre
 pluripotente, 119
 totipotente, 188
- célula somática, 97
 ciclo celular, 100
 híbrida, 221
 híbrida, muestras para preparar,
 222

- célula
competente, 211
en cultivo, como muestra, 121
híbrido monocromosómico, 224
híbrido subcromosómico, 224
métodos de separación, 122
recambio, 415
vida media. Véase apoptosis
- centimorgan, 220
- centrifugación
de equilibrio de sedimentación, 124
diferencial, 122, 133
fundamento, 122
isopícnica, 67, 113, 124, 137, 143
para elutriación, 125
ultracentrifugación, 136
zonal, 124, 136
- centro
de ajuste, 280
de ramificación, 280
- centrómero, 87, 88, 92, 98
en vectores de clonación, 211
y DNA satélite, 113
- CFTR. Véase regulador de la conductancia transmembrana
- cGMP. Véase GMP cíclico
- Chambord (castillo de), 47
- chaperona. Véase carabina
- chaperonina. Véase carabina
- Chargaff (reglas de), 38, 42
- chip de DNA. Véase biochip
- choque térmico. Véase carabina
- ciclina, 424
A, 424
B, 424
complejo con Cdk, 424, 426, 427
D, 417, 424
de G₁, 423, 424
de G₂, 423, 424
de inicio. Véase ciclina de G₁
degradación, 426, 428
E, 424
F, 424
mitótica. Véase ciclina de G₁
síntesis, 424, 426, 428
- ciclo
de elongación en la traducción, 314
del factor eIF-2, 310, 320
del factor eEF-1 α , 312
- ciclo celular, 97
estímulos de detención, 429, 431
estímulos de progresión, 428
fases, 98, 100
freno por p53, 431
oncogenes relacionados, 429
puntos de control, 422
regulación, 422, 427
y condensación, 100
- ciclofosfamida, 355
- cicloheximida, 316
- cigoto, 100, 103
- cinetocoro, 92
- CIP1 (proteína), 426
- cirrosis, 376
- cis
elemento, factor o secuencia, 264
Véase también Golgi
- cisplatino, 355, 359
- cisteína, en centro activo
de la alquiltransferasa, 358
de las caspasas, 432
- cisteína endopeptidasa, 432
- cistrón, 301
- citidilato, 27
- citidina, 20
- citocina, 437
- citocinesis, 99
- citocromo c, 328
como señal de activación de apoptosis, 433
- citofluorímetro, 126
- citogenética, 88, 91
- citómetro de flujo, 126, 225
- citoquinesis. Véase citocinesis
- citosina, 12
metilada, 272
- CKI. Véase Cdk, inhibidores
- clasificación de proteínas.
Véase tráfico
- clon
biblioteca de, 215
concepto, 187
neoplásico, 413
- clonación
introducción general, 187
objetivos, 203
- clonación acelular. Véase PCR
- clonación celular, 188, 197
esquema general, 204
etapas, 204
para obtención de sondas, 176
preparación del inserto, 205
- clonación funcional, 380
- cloranfenicol, 316
- cloroformo
en la extracción de DNA, 134
- cloroplastos, 5, 323
- cloruro cálcico, 211
- cloruro de cesio. Véase centrifugación isopícnica
- cM. Véase centimorgan
- CMP-sialico, 332
- coactivador, 270
- cobalamina, 22
- Cockayne (síndrome de), 363
- codificación
relación de, 288
unidad de, 288
- codificante. Véase DNA codificante
- código genético, 287
degeneración, 295
diferencias, 298
mitocondrial humano, 298
representación circular, 294
- código genético
representaciones en tabla, 293, 295
segundo código, 304
universalidad, 298
- codominante, 95, 394
- codón, 288, 289
apareamiento con el anticodón, 296
asignación de aminoácido, 290
AUG o de inicio, 294, 308
de paro o terminación, 294, 314
número necesario, 288
número por aminoácido, 295
sin sentido.
Véase codón de paro
sinónimos, 295
solapantes, 288
superpuestos, 288
- coeficiente de extinción.
Véase absorptividad
- coeficiente de sedimentación, 80
- coenzima A, 30
- coenzimas
nucleosídicas, 22
nucleotídicas, 30
- cofactor general, 270
- cola de poli(A), 139, 206, 280
efecto protector, 286
- colágeno, 328, 329, 414
- colchicina, 90
- colonias. Véase bacteria
y véase clon
azules, método de, 214
- compactación.
Véase condensación
- compartimentos subcelulares, 322
- compatibles, 202
- competente, 211
- complejidad del genoma, 107, 109
estudio de. Véase Cot
- complejo
abierto, 253
binario en la traducción, 307
de empalme.
Véase ayustosoma
de iniciación 80S, 309, 310
de inicio de la transcripción, 267
de inicio de la transcripción, modelos, 267
de preiniciación 40S, 309
de reconocimiento del origen, 151
de reparación, 360
de translocación, 325
proteína multicatalítica, 342
ternario en la traducción, 307
- complementariedad, 39
imperfecta, 169
perfecta, 169
- condensación del DNA, 65, 83, 91, 98
en la apoptosis, 432
grado de, 85
niveles, 85, 100

condensación del DNA,
relación con la transcripción, 250
y acetilación de histonas, 264
y ciclo celular, 100
y regulación de la transcripción,
263

conductora (hebra), 152-154

conformación
anti y sin de los nucleósidos en
el DNA, 55
C2'-endo y C3'-endo, 54, 69
de la desoxirribosa
en A- y B-DNA, 54
relajada, 66
superenrollada, 67

conservadora. Véase mutación
por sustitución
y véase replicación

constitutivo (gen), 269

constricción secundaria, 88

cóntigo, 229

controles G/S, G/M y M, 422,
423

cordocentesis, 120

corea de Huntington.
Véase Huntington

corion, 120

corrección de errores
en la traducción, 305

corrección de pruebas, 147

corriente abajo, 249

corriente arriba, 249

cortadora infrecuente, 216, 226,
229

corte y empalme. Véase ajuste
corte y pegado. Véase ajuste
corte tisular, hibridación, 172

cos, 210

cósmido, 210

Cot, 116

cotranscripción, 253

cotransformación.
Véase transformación

CPRG, 186

CPSF, 279

cremallera de leucina, 64

creatinquinasa, 437

Creutzfeld-Jacob (enfermedad de)
relación con el plegamiento, 337

Cro (proteína del fago lambda), 62

crómátida, 85, 88, 98
separación anómala, 410
y ciclo celular, 100

crómatina, 5, 84
actividad transcripcional, 87,
263
condensación, 98
descondensación, 98
extendida, mapeo por FISH,
225
remodelado, 263

cromo, 355

cromosoma, 5
acéntrico, 92, 408, 409
acrocéntrico, 88, 409

cromosoma
anomalías. Véase
cromosomopatías

anular, 408

artificial, 211

asas, 85, 87

brazos, 88, 99

clasificación, 88

derivado, 409

dicéntrico, 409

hibridación, 172

homólogos, 93

interfásico, 85, 99

interfásico, mapeo FISH, 225

isocromosomas, 410

metacéntrico, 88

metafásico, 85, 88, 99

metafásico, mapeo FISH, 225

morfología, 88

número, 6, 7, 93

pares, 93

pintado de, 172

recombinantes, 218

reparto en la división celular,
105

rotura, 406

satélite cromosómico, 88

segregación, 104, 105

sexual, 93, 94

submetacéntrico, 88

tamaño, 88

X, 93

X, inactivación, 87, 264

Y, 93

cromosomopatía, 392
adquirida, 402
concepto, 401
constitutiva, 402
estructural, 406
nomenclatura, 402
numérica, 403
somática, 402

CRP. Véase proteína receptora de
AMP cíclico

cruciforme, 59

CSA, 363

CSB, 363

CTD, 252, 253, 279
asociación con la caperuza 5',
279
en inicio y elongación de la
transcripción, 254, 255
fosforilación, 254, 255, 267

CTF, 269

cuerpo
apoptótico, 433
de empalme.
Véase ayustosoma
de inclusión, 339
polar, 103

cumermicina, 259

CYFRA 21.1, 436

D

d4T. Véase estavudina

dactinomicina, 355

daltonismo, 397

DAPI, 91

daunomicina, 259

daunorrubicina, 259

ddc (gen), 418

ddC. Véase zalcitabina

ddl. Véase didanosina

ddNTP. Véase didesoxinucleótido

DEAE-dextrano, 212

dedo de zinc, 63, 265

defectos nutricionales, para la
selección de clones
recombinantes, 214

degeneración
aplicación al diseño de sondas,
180
del código genético, 295
e impacto de las mutaciones, 295

deleción, 347, 348.
Véase también mutación
cromosómica intersticial, 408
cromosómica terminal, 406

densidad. Véase gradiente
desacetilasa de histonas, 264

desaminación, 353
reparación, 360, 362

desarrollo embrionario, 285

descondensación, 84, 85, 98.
Véase también
condensación

desenrollamiento, 151, 248
en la transcripción, 253-254

desequilibrado, 406

desfosforilación, 327
del extremo 5' del mRNA, 278

desmetilasa, 273
en regulación de la expresión, 272

desnaturalización, 163
agentes causantes, 164
curva de, 166
del promotor, 253
en la PCR, 188, 189
local por la unión de proteínas,
61
reversible. Véase
renaturalización
y contenido en G+C, 166
y propiedades del DNA, 165

desoxiribofosfodiesterasa, 361

desoxirribonucleasa.
Véase DNasa

desoxirribosa, 11

desplazamiento de la mella. Véase
traslado de la mella

despurinación, 354

destino
de proteínas, 322

detección
de clones recombinantes, 213
directa e indirecta, 180

- dextrógiro, 45
 diagnóstico prenatal, 120, 386
 diana, 168
 de restricción, 199
 dicéntrico. Véase cromosoma
 dictiosoma, 323
 didanosina, 23
 didesoxinucleótido, 235
 dietilestilbestrol, 355
 diferenciación, 285
 difosfato, 10
 difracción de rayos X
 para la estructura del DNA, 37
 para la estructura del tRNA, 75
 DIG. Véase digoxigenina
 digestión
 doble, 227
 enzimática, 122
 parcial del DNA genómico, 216
 digoxigenina, 186
 dihidrouridina, 72
 dímero de timina, 356
 reparación, 359
 dimetilnitrosamina, 355
 dimetilsulfato, 355
 para secuenciación, 232
 dimetoxitritilo, 179
 dinucleósido polifosfato, 29
 dioxetano, 186
 dioxigenasa, 328
 diploidia, 6, 93, 100, 403
 disenzimias, 392
 disgregación
 de tejidos, 133
 química, 122
 disomía, 403
 displasia, 413
 distal. Véase promotor
 distrofia muscular de Duchenne, 397
 disyunción, 104
 defecto de, 404, 405
 diversidad genética.
 Véase polimorfismo
 división celular
 meiótica. Véase meiosis
 mitótica. Véase mitosis
 reductora. Véase meiosis
 DMT, 179
 DNA, 1
 7S, 160
 agrupado, 108, 109
 altamente repetitivo, 113
 análisis de tamaño por
 electroforesis, 137, 138
 basura, 109
 chatarra, 109
 circular, 145
 codificante, 8, 108, 243, 245,
 248, 250
 comparación entre las formas A,
 B y Z, 52
 complementario. Véase cDNA
 composición. Véase G+C
 daño y proteína p53, 431
 de copia única, 8
 DNA
 densidad, 137
 disperso, 108, 109
 escalera, 137, 138, 433
 espaciador, 85, 86, 263, 433
 esqueleto, 84
 fetal, 120
 forma A. Véase A-DNA
 forma B. Véase B-DNA
 forma mutante, 344
 forma salvaje, 344
 forma silvestre, 344
 forma Z. Véase Z-DNA
 fragmentación en la apoptosis,
 432, 433
 fragmentación, análisis
 electroforético, 433
 hijo, 141
 iniciador, 154
 interacción con proteínas, 60
 intergénico, proporción en
 el genoma, 108
 internucleosómico.
 Véase DNA espaciador
 ligero, 143
 mapeo por FISH, 225
 marcage, 143
 masa en la célula, 6
 metabolismo, 141
 metilación, 273
 microsatélite, 114, 383
 microsatélite, aplicación a la
 huella genética, 385
 minisatélite, 114, 383
 minisatélite, aplicación a la
 huella genética, 385
 minisatélite hipervariable, 384
 mitochondrial, 108, 257
 mitochondrial, utilidad, 377
 mitochondrial, muestras, 377
 moderadamente repetitivo, 113
 no codificante, 8, 108, 245, 250
 pesado, 143
 progenitor, 141
 proporción de bases, 38
 reasociación. Véase
 renaturalización
 recombinante. Véase rDNA
 reparación. Véase reparación
 del DNA
 repetitivo, 8, 108, 109
 repetitivo agrupado, 109
 repetitivo codificante, 110
 repetitivo codificante disperso,
 112
 repetitivo disperso, 109
 repetitivo no codificante, 113, 383
 repetitivo, análisis, 116
 satélite, 87, 113, 137, 383
 tipos en el genoma, 108
 variantes estructurales, 51
 velocidad de síntesis, 149
 DNA glicosilasa. Véase
 N-glicosilasas de DNA
 DNA ligasa. Véase ligasa
 DNA metiltransferasa.
 Véase metilasa
 DNA polimerasa, 84, 147
 actividades enzimáticas, 148
 cambio de α a δ o ϵ , 154
 centros activos, 147
 del fago T7, 235
 empleo para el marcaje de
 sondas, 182
 en la reparación, 359, 361
 eucarióticas, 148
 fragmento de Klenow, 147
 masa molecular, 149
 Pfu, 193
 procarióticas, 147-148
 propiedades, 148
 subunidades, 147, 149
 Taq, 189, 235
 termoestable, 189
 tipos, 148
 Tth, 196
 ubicación subcelular, 148
 velocidad, 148
 I, 147-148
 I, aplicación para la preparación
 de sondas, 182
 I, en la síntesis de cDNA, 206
 I, fragmento de Klenow, 183,
 184, 235
 II, 147-148
 III, 147-148
 III, modelo molecular, 149
 α , 148, 152-154
 α , en la síntesis de telómeros,
 160
 β , 148
 β , actividad fosfodiesterasa, 361
 δ , 148, 153-154
 ϵ , 148, 153-154
 γ , 148
 η , 148
 ι , 148
 ζ , 148
 DnaA, 151
 DnaB, 151
 DnaC, 151
 DNAPol. Véase DNA polimerasa
 DNasa, 85, 199
 I pancreática, 182
 II, 199
 doble hebra, 41
 doble hélice, 45
 apilamiento, 48
 carácter hidrofílico, 47
 carácter hidrofóbico, 48
 curvatura, 57
 de tipo A en el DNA.
 Véase A-DNA
 de tipo A en el RNA, 69
 enlaces de hidrógeno, 48
 formación de triple hélice, 60
 fuerzas estabilizadoras, 48
 parámetros, 46
 posición de los fosfatos, 47
 relación con la replicación, 49

dogma central de la genética molecular, 2, 300
comparado en citosol y mitocondria, 243
dolicol, 332, 333
fosfato, 333
pifosfato, 333
doméstico (gen), 269
dominante, 95, 394
dot-blot, 170
Down (síndrome de), 406
doxorubicina, 259
dRPasa. Véase desoxirribonucleasa
Duarte. Véase galactosemia
Duchenne (distrofia muscular de), 397
duplicación, 367
dUTP pifosfatasa, 362

E

E1A (proteína), 431
E1B (proteína), 431
E2F (proteína), 430
E6 (proteína), 431
E7 (proteína), 431
edad materna
y riesgo de trisomía, 406
Edward (síndrome de), 406
eEF. Véase factor de elongación de la traducción
efecto hidrofóbico, 338
EGF. Véase factor de crecimiento epidérmico
eIF. Véase factor de iniciación de la traducción
elastasa
inhibición por antitripsina, 376
elastina, 376, 414
electroforesis, 171, 173
calibración, 138
en campo pulsante, 138, 226
fundamento, 137
preparación de geles, 137
tipos de gel y tamaño de DNA, 137
electroporación, 212
electrostática (interacción), 185
elemento
de respuesta. Véase promotores distales
de respuesta a hierro, 319
de respuesta a hormonas, 270
específico de regulación. Véase promotor distal
móvil. Véase transposón nuclear disperso corto.
Véase SINE
nuclear disperso largo.
Véase LINE
transponible. Véase transposón
ELL, 255
elongación de la replicación, 152
etapas, 153
mecanismo, 153, 154

elongación de la traducción
esquema general, 312
translocación, 313
transpeptidación, 313
ubicación, 311
elongación de la transcripción
complejo de, 254
factores, 255
elongación en la PCR, 189
elutriación centrífuga, 125
embriogénesis
metilación del DNA, 272
empaquetamiento, 83
endocitosis, 323
endomitosis, 94
endonucleasa, 198
AP, 361
de restricción. Véase enzimas de restricción
de tipos 'a' y 'b', 198
en la reparación, 361
FEN1. Véase FEN1
generación del extremo 3' del mRNA, 279
endosoma, 323
enfermedad. Véase también bajo el nombre específico
asociada a la reparación del DNA, 363
autosómica, concepto, 392, 394
autosómica dominante, 395
autosómica recesiva, 395
búsqueda del gen responsable, 221, 385
citogenética.
Véase cromosomopatía
clasificación molecular, 390
compleja, 399
congénita, diagnóstico, 386, 387
cromosoma afectado, 392
cromosómica.
Véase cromosomopatía
genética, 369, 379, 390
hereditaria, 221, 345
ligada al sexo, 392, 394
ligada a X dominante, 396
ligada a X recesiva, 397
ligada a Y, 397
mitocondrial, concepto, 392
mitocondrial, 398
mixta, 399
molecular, 389
monogénica, 392, 393
multifactorial, 399
poligénica, 399
productos génicos, 392
enfisema, 376, 396
relación con el plegamiento, 337
enlace
de hidrógeno de tipo Hoogsteen, 60, 73, 296, 297
de hidrógeno de tipo Watson y Crick, 296, 297
disulfuro, 335
fosfoanhidrido, 10
fosfodiéster, 11, 146, 247

enlace
fosfodiéster atípico [5'-2'], 281
fosfoéster, 9
N-glicosídico, 20, 329
N-glicosídico, hidrólisis, 354, 361
O-glicosídico, 329
S-glicosídico, 330
isopeptídico, 341
peptídico, formación, 313
tioéster con ubiquitina, 341
topológico, 68
trifosfato [5'-5'], 278
enrollamiento
en la transcripción, 254
ensamblado, 240
enteroquinasa, 336
envejecimiento
y telómeros, 160
enzima
activadora de ubiquitina, 341
como marcador de sondas, 180
fotorreactivadora, 357
hidrolítica en lisosomas, 341
portadora de ubiquitina, 341
enzima de restricción, 184, 197
aplicación al mapeo, 226, 227
características generales, 199
como mecanismo de defensa en las bacterias, 200
cortadora infrecuente.
Véase cortadora
de tipo I, 199
de tipo II, 199, 201
de tipo II, bacteria de origen, 202
de tipo II, ejemplos concretos, 202
de tipo II, nomenclatura, 202
de tipo II, secuencias diana, 202
de tipo III, 199
en la preparación de genotecas, 216
especificidad, 199
estructura, 199
frecuencia de reconocimiento, 226
nomenclatura, 199
para linealizar un plásmido, 177
para síntesis de sondas, 184
sitios de reconocimiento, 199
epigenético, 271, 412
equilibradas
hebras del DNA, 44
equilibrado, 406
equivalente genómico, 216
erA, 417
erB, 417
ERCC1, 360, 363
eRF. Véase factor de terminación de la traducción
eritromicina, 316
error congénito del metabolismo, 390
escalonado (extremo), 201
escintileación, 359, 360

escisión
 de bases nitrogenadas, 358, 360
 de nucleótidos, 358, 359
 escorbuto, 328
 esparsomicina, 317
 especie reactiva de oxígeno, 354, 377
 reparación, 360
 espectrofotometría, 186
 espermátida, 102
 espermatozoides, 102
 espermatogénesis, 102
 espermatogonio, 102
 espermatozoide, 84, 102
 espermidina, 34
 espermina, 34
 esqueleto nuclear, 84, 87
 EST, mapa físico, 228
 estabilidad del RNA, 286
 estavudina, 23
 estearato, 334
 esterasa y viabilidad, 131
 estimulador, 266
 estreptavidina, 129, 186
 estreptolidigina, 258
 estreptomycin, 316
 estructura
 tridimensional de proteínas, predicción, 338
 terciaria, relación con la secuencia, 338
 estructural. Véase gen, secuencia
 estudio familiar, 385
 etidio. Véase bromuro de etidio
 etilmetanosulfonato, 355
 etiqueta. Véase péptido señal
 etiqueta de secuencia expresada. Véase EST
 eucariote. Véase eucariota
 eucariota, 1
 diferencias con procarionte, 5
 eucromatina, 85, 87
 y actividad transcripcional, 250
 euploidia, 403
 exclusión
 ensayos de, 129
 exocitosis, 323
 exón, 277
 concepto, 109, 245, 276
 disposición en el gen, 245
 proporción en el genoma, 108
 exonucleasa, 198
 3' y 5', 147
 5' en marcaje de sondas, 182
 expresión diferencial, 269
 expresión génica, 2
 basal, 269
 diferencial, 261
 niveles de regulación, 262
 regulación, 261, 269
 regulación postranscripcional, 284
 extracto S-30, 289

extremo
 3', modificación, 279
 5', modificación, 277
 cohesivo, 201
 cohesivo, asociación, 203
 cohesivo, en telómero, 158
 compatible, 202
 romo. Véase romo

F

FACS, 127, 225
 factor
 activador de las proteasas apoptóticas, 433
 asociado a TBP, 267, 270, 271
 de coagulación VIII, 397
 de coagulación, carboxilación, 327
 de crecimiento, 416
 de crecimiento derivado de plaquetas, 416, 421
 de crecimiento epidérmico, 417, 421
 β de crecimiento transformante, 418
 de replicación C, 154
 de terminación mitocondrial, 257
 estimulador de la poliadenilación por escisión, 279
 inhibidor del crecimiento, 418
 promotor de la maduración o de la fase M, 424
 factores de elongación de la traducción, 311
 eEF-1 α , 311, 312
 eEF-1 β , 311, 312
 eEF-2, 312, 313
 factores de iniciación de la traducción, 309
 eIF-2, 309, 310
 eIF-2B, 310
 eIF-3, 309
 eIF-4A, 309, 310
 eIF-4B, 309, 310
 eIF-4C, 309
 eIF-4E, 309, 310, 319
 eIF-5, 309, 310
 eIF-6, 309
 factores de terminación de la traducción
 eRF, 314
 factores de transcripción, 60, 84, 253
 asociación en el complejo de inicio, 270
 coactivadores, 270
 cofactores generales, 270
 como oncoproteínas, 417
 como oncosupresores, 418
 de corriente arriba, 269
 de la elongación, 255
 definición, 264
 E2F, 430
 función, 264
 generales, 267

factores de transcripción
 inducibles, 269
 interacción con la RNA polimerasa, 270
 mitocondrial, 257
 motivos estructurales, 265
 Myc, 430
 proximales, 269
 p53, 431
 SREBP, 432
 TFI, 265, 271
 TFII, 265, 266
 TFIIA, 268
 TFIIB, 268
 TFIID, 267, 268
 TFIIIE, 253, 268
 TFIIF, 253, 268
 TFIH, 253, 268, 363
 TFIH, como helicasa, 267
 TFIH, como proteína quinasa, 267
 TFIH, en la reparación, 359, 360
 TFIH, subunidades, 360
 TFII-L, 270
 TFIII, 265, 271
 TFIIIA, 63, 271
 TFIIB, 271
 TFIIC, 271
 FAD / FADH₂, 30
 coenzima de la fotoliasa, 357
 fago. Véase bacteriófago
 falciforme
 detección por RFLP, 382
 herencia, 396
 mutación causante, 349
 polimorfismo, 382
 Falconi (anemia de), 363, 406
 FAM, 238
 familia multigénica, 110, 367
 farnesilo, 335
 farnesiltransferasa
 y terapia antitumoral, 335
 fase
 analítica, 118
 de reposo, 98
 G₀, G₁, G₂, S y M, 98
 postanalítica, 118
 preanalítica, 118
 Fast Red TR, 186
 FCR. Véase fuerza centrífuga relativa
 fecundación, 100, 106
 FEN1, 156
 en la reparación, 361
 en la síntesis de telómeros, 160
 fenilalanina-4-monooxigenasa, 396
 fenilcetonuria, 396
 fenol
 en la extracción de DNA, 134
 fenotipo, 95
 fenoxazona, 258
 ferritina, 319
 fetal
 DNA, 120
 sangre, 120

α -fetoproteína, 435
 fetoscopia, 120
 fibra
 básica de cromatina.
 Véase fibra de 10 nm
 de 10 nm, 85, 86, 87
 de 10 nm y control
 pretranscripcional, 263
 de 10 nm y transcripción, 250
 de 30 nm, 85, 87
 fibrosis quística
 herencia, 396
 mutación causante, 351
 relación con el plegamiento, 337
 Ficoll, 124
 FISH, 91, 172
 cromosómica, 224
 de alta resolución, 225
 fitohemaglutinina, 90
 fluctuación. Véase balanceo
 fluoresceína, 131, 182, 238
 fluorimetría, 186
 fundamento, 126
 fluorocromo, 126, 172, 186, 225,
 232
 en marcaje de sondas, 180, 182
 en secuenciación del DNA, 238
 fluoróforo, 126
 fMet. Véase N-formilmetionina
 FMN, 30
 formaldehído, 165
 formamida, 165, 169
 formamidopirimidina-DNA
 glicosilasa, 362
 formazano, 131
 formiato, para secuenciación, 232
 formiato-tetrahidrofolato ligasa,
 308
 fórmico (ácido), para
 secuenciación, 232
 N-formilmetionina, 308
 formiltetrahidrofolato, 308
 fos (gen), 417
 fosfatasa ácida prostática, 436
 fosfatasa alcalina, 184, 186, 208
 isoenzimas por duplicación
 génica, 367
 fosfatasa inorgánica, 303
 fosfato cálcico, 212
 fosfato inorgánico (P_i), 9
 fosfato α en NTP.
 Véase nucleofílico
 y véase marcaje de NTP
 fosfato γ en NTP.
 Véase marcaje de NTP
 fosfodiesterasa
 I, de veneno de serpiente, 198
 II, de bazo bovino, 198
 fosfoglucomutasa, 374
 fosfoproteína fosfatasa, 425
 fosforamida, 179
 fosforilación
 como mecanismo regulador, 327
 de aminoácidos en las
 proteínas, 327

fosforilación
 de eIF-2, 320
 del dominio CTD, 252, 254, 255
 reversible, 327, 425
 fosforilasa fosfatasa, 327
 fosforilasa quinasa, 327
 fosfoserina/fosfotreonina
 fosfatasa, 327, 425
 fosfotirosina fosfatasa, 327, 425
 fotodímero, reparación, 357
 fotolisa, 357
 fotoproducto, 356
 fracción de recombinación.
 Véase frecuencia
 fracciones subcelulares, 133
 fragmento de restricción, 201
 frecuencia de recombinación, 219, 220
 fucosa, 330
 fucosiltransferasa, 372
 fuerza centrífuga relativa, 122
 funiculocentesis, 120
 fusídico (ácido), 317
 fusión celular, 222
 fusión del DNA. Véase
 desnaturalización
 FIIF, 255

G

G+C (contenido), 39, 137, 166
 en regiones terminadoras de
 transcripción, 255
 GADD, 431
 galactoquinasa, 374
 galactosa, 330
 vías metabólicas, 374
 galactosa-1-fosfato
 uridiltransferasa, 373
 galactosaminiltransferasa, 372
 galactosemia, 373-375
 β -galactosidasa, 186.
 Véase también lacZ
 galactosiltransferasa, 372
 GALT. Véase galactosemia
 gametogénesis, 99
 gameto, 94, 97, 99
 ganciclovir, 13
 GC (secuencia o caja), 266, 269
 GDP-fucosa, 332
 GDP-manosa, 332
 gel. Véase electroforesis,
 agarosa y poliácridamida
 gemelos, 386
 gen
 análisis, 380
 búsqueda. Véase gen, análisis
 comparado con cistron, 301
 con varios productos génicos,
 246, 276
 constitutivo, 269
 de clases I, II y III, 265
 de expresión diferencial, 269
 definición, 244-246, 276
 doméstico o casero, 269
 eucariótico, estructura, 245

gen
 expresión.
 Véase expresión génica
 fragmentos de, 112
 fragmentos de, y transcripción,
 250
 inducible, 266, 269
 marcador, 214
 monomorfo, 368
 nomenclatura, 246
 oncosupresor.
 Véase oncosupresor
 organización, 277
 polimorfo, 368
 región no traducida 5' y 3', 245,
 276
 región no traducida interna.
 Véase intrón
 región traducida, 245
 región transcrita, 245
 región transcrita eliminada antes
 de la traducción, 245
 regulable, 269
 secuencia estructural, 109, 245
 secuencia reguladora, 109, 245
 secuencia terminadora, 245
 solapante, 246, 249
 supresor de tumores.
 Véase oncosupresor
 truncado, 112
 variantes, 112, 369
 genealogía, 385
 genochip. Véase biochip
 genoma, 1, 5
 codificante, 250
 datos, 6
 diversidad. Véase polimorfismo
 extranuclear, 5
 fracción transcrita, 250
 mitocondrial humano, 8
 nuclear, 5
 número de cromosomas, 6
 número de genes, 8, 241
 tamaño, 6, 7
 tipos de DNA, 108
 genómica, 217, 390
 comparativa, 241
 estructural, 379
 genosensor. Véase biochip
 genoteca, 215
 de cDNA, 216
 de cDNA, en la búsqueda de
 genes, 380
 genómica, 216
 número de clones, 216
 genotipo, 95
 gentamicina, 317
 geometría de los pares de bases
 comparada en A- y B-DNA, 54
 geranilgeranilo, 335
 geranilo, 335
 Giemsa (tinción), 87, 91
 girasa, 152
 inhibición por antibióticos, 259
 glicano. Véase oligosacárido

- N-glicano, 330
 estructura central, 330
 reacciones de biosíntesis, 333
 tipos complejo, híbrido y rico en manosa, 331
 O-glicano, 330
 reacciones de biosíntesis, 332
 glicolípido, en la biosíntesis de N-glicanos, 333
 glicoproteínas. Véase glicosilación
 glicosaminoglicano, 329
 glicosilación, 328
 en Golgi, 332, 334
 en N-, 330
 en O-, 330
 en retículo endoplásmico, 333
 localización subcelular, 332
 propiedades que induce en la proteína, 328
 tipos, 329
 N-glicosilasa de DNA, 361, 362
 glicosiltransferasa, 332
 A y B de grupo sanguíneo ABO, 370, 371
 globina
 expresión en el desarrollo animal, 112
 familia multigénica, 111
 herencia, 396
 origen evolutivo, 111
 polimorfismo, 382
 tipos, 112
 velocidad de traducción, 320
 y talasemia, 396
 β^0 , 349, 382
 glóbulo fundido, 338
 gluco-. Véase glico-
 glucógeno fosforilasa, 327
 glucosa, 330
 γ -glutamyltranspeptidasa
 como marcador tumoral, 437
 glutaraldehído, 185
 glutatión, 335
 GMP cíclico, 29
 Golgi (complejo de), 323
 gonadotropina coriónica, 436
 Gppp, 278
 gradiente de densidad, 124, 137, 143
 gránulos de secreción, 337
 Grisolia, 239
 GroEL, 339
 GroES, 339
 Grunberg-Manago y Ochoa, 246
 grupo sanguíneo ABO
 anticuerpos, 369
 antígenos, 369
 bioquímica, 372
 ensayo serológico, 370
 fenotipos, 369, 373
 genética, 370
 genotipos, 373
 localización cromosómica, 371
 puntos polimórficos, 371
 resumen, 373
 grupo prostético, 336
 GTBP, 362
 GTPasa
 proteína Ras, 421
 guanidina, 134
 guanina, 12
 guanilato, 27
 guanililación
 del extremo 5' del mRNA, 278
 guanosilación
 del extremo 5' del mRNA, 278
 guanosina, 20
 guía (hebra). Véase conductora
 Guthrie (tarjeta), 120
- ## H
- H. Véase hebra H
 H (gen), 370
 haplogrupo, 377
 haploidia, 6, 93, 100, 403
 haplotipo, 220
 HAT (medio de cultivo), 223
 hCG, 436
 HCl, 320
 H-DNA, 60
 hebra
 antisentido, 248
 codificante, 248
 con sentido, 248
 H. Véase DNA mitocondrial informativa, 248
 L. Véase DNA mitocondrial ligera, 45, 160
 molde, 248.
 Véase también molde
 negativa, 248
 no codificante, 248
 no informativa, 248
 no molde, 248
 no transcrita, 248
 pesada, 45, 160
 positiva, 248
 transcrita, 248
 helicasa, 150, 151
 de híbrido RNA-DNA, 255
 en la reparación, 359
 en la terminación de la transcripción, 255
 en la transcripción, 253
 en TFIIH, 267
 hélice de reconocimiento, 61
 hélice-bucle-hélice, 62
 hélice-giro-hélice, 61
 hélice-vuelta-hélice, 61
 helicidad del DNA, 45
 hemo, 336
 como regulador de la traducción, 320
 hemofilia A, 397
 hemoglobina
 falciforme o S, 349, 382
 glicosilada, 329
 hemoglobinopatía, 349
 herencia, 2. Véase también árbol genealógico
 de las mutaciones, 344
 herencia
 del cáncer, 434
 materna, 398
 patrón mendeliano, en enfermedades monogénicas, 393
 patrón no mendeliano, 398
 Hershey y Chase, 2
 heterocariote, 222
 heterocigosis
 y polimorfismo, 366
 heterocigótico, 95
 heterocromatina, 85, 87, 88, 113
 constitutiva, 87, 92
 facultativa, 87
 heterodúplex, 163, 168
 heterogamético, 94
 heteroplasma, 398
 heteroploidia, 403
 hexoquinasa, 374
 hexosa fosfato isomerasa
 como marcador tumoral, 437
 hexosaminidasa A
 herencia, 396
 mutación, 350
 HGPRT, 223
 deficiencia, 224
 HHR23B, 360
 hibridación, 163
 análisis molecular, 167
 de DNA, 171
 de RNA, 171
 de Southern. Véase Southern
 en dot-blot, 170
 en fase líquida, 169
 en la PCR, 189
 en slot-blot, 170
 en soporte sólido, 170. Véase también sobre biochips
 factores que afectan, 168
 fundamento, 168
 in situ, 172
 in situ por fluorescencia.
 Véase FISH
 métodos, 169
 Northern. Véase Northern
 para la detección de clones recombinantes, 213
 rigor, 168
 sobre biochips, 388
 sobre cortes tisulares, 172
 sobre preparaciones cromosómicas, 172
 híbrido
 DNA-DNA, 163, 167
 DNA-RNA, 69, 163, 249, 254
 DNA-RNA, desestabilización en la terminación de transcripción, 256
 RNA-cDNA, 196, 206
 RNA-RNA, 163
 hidrato de pirimidina
 DNA glicosilasa, 362
 hidrazina para secuenciación, 232
 hidrocarburos aromáticos
 policíclicos, 355

- hidrofobicidad, 14
 influencia en el plegamiento, 338
- hidrólisis
 alcalina del RNA, mecanismo, 36
 química de ácidos nucleicos, 35, 354
- hidroxiapatito, 139, 169
- hidróxido (ion), 354
- 8-hidroxiguanina, 354, 356
- hidroxilación
 de aminoácidos, 328
- hidroxilasa, 328
- hidroxilo (radical), 354, 356
- hidroxiprolina, 328
- hidruro de níquel, 311
- hierro, metabolismo, 319
- hipercolesterolemia familiar, 395
- hipercrómico, 165
- hiperplasia, 413
- hipervariable. Véase regiones y véase minisatélites
- hipocrómico, 165
- hipótesis
 un gen-un polipéptido, 244
 un gen-un producto génico, 245
 un gen-un RNA, 245
 un gen-una enzima, 244
 un gen-una proteína, 244
- hipoxantina, 21
 como resultado de la desaminación, 353
 en medio HAT, 223
- hipoxantina/guanina
 fosforibosiltransferasa. Véase HGPRT
- histona, 34, 84, 86
 acetilación, 264, 326
 familia multigénica, 110
 H1, 84, 86, 87
 H2A, 84, 86
 H2B, 84, 86
 H3, 84, 86
 H4, 84, 86
 H4, acetilación, 326
 H4, metilación, 328
 H5, 84, 86
 reducción de la carga positiva, 264
 tipos, 84
- histona acetiltransferasa, 264
- histona desacetilasa, 264
- HLH. Véase hélice-bucle-hélice
- HMG, 84, 86
- hMLH1, 362, 363
- hMSH2, 362, 363
- hMSH6, 362
- HNPCC, 363
- HNPP, 186
- hnRNA, 70, 251
- Hoechst, 91, 126
- Hogness (caja de). Véase TATA
- holoenzima. Véase modelo homeodominio, 62, 265
- homocariote, 222
- homocigótico, 95
- homodúplex, 163, 168
- homogamético, 94
- homogeneización, 133
- homología de secuencia e hibridación, 167
- homopolinucleótido, 289
- homopolipéptido, 290
- Hoogsteen. Véase enlaces de hidrógeno de tipo Hoogsteen
- hormonas esteroides
 receptores de, 270
- hormonas tiroideas
 receptores de, 270
- horquilla, 59, 70
 de replicación, 145, 150, 151, 155.
 Véase retardada de RNA, 319
 y terminación de la transcripción, 255
- hospedadora.
 Véase célula anfitriona
- hPMS, 362
- HRE. Véase elementos de respuesta a hormonas
- Hsp. Véase proteínas de choque térmico
- HTH. Véase hélice-giro-hélice
- hTR. Véase telomerasa, componentes
- huella dactilar genética, 377, 385
- huella de DNA
 y DNA minisatélite, 114
- humo, productos mutágenos, 355, 359
- Huntington (enfermedad de), 395
- huso mitótico, 99
- I**
- IBP. Véase proteínas de unión a Irr
- ICR, 271
- identidad, pruebas de, 385
- idiograma, 89
- IF. Véase factores de iniciación de la traducción
- imina-amina. Véase tautomería
- inactivación insercional, 214
- indicador, 181, 185
- individualidad genética, 366
- inducible (gen), 269
- información genética, 1, 243
- informativa, 248
- inhibidor controlado por hemo, 320
- inhibidor (promotor, elemento o secuencia), 266
- iniciación al azar, 183
- inicio (punto de control del ciclo celular), 423
- inicio de la replicación, 150
 proteínas de, 150, 151
- inicio de la traducción, 307
 energética, 310
 etapas, 309
- inicio de la traducción por el extremo N, 300
- inicio de la transcripción
 como punto de control, 264
 complejo de, 253, 267
 determinación del punto de inicio, 267
 modelos, 253
 modelo del holoenzima, 267
 modelo escalonado, 267, 268
 transición a la elongación, 254, 267
- INK4, 426
- inmunoanálisis, 128
- inmunoglobulina M
 gen y ajuste alternativo, 286
- inosina, 21, 72
 en el anticodón, 296
 en el balanceo, 297
- Irr, 253, 265, 270
- insecto, 210
- inserción, 347, 348
 método de preparación de virus recombinantes, 210
- inserto, 203, 204
 de cDNA, 205
- insulina, 337
- integrina, 413, 414
- intensificador, 266
- interacción
 de tRNA con proteína, 305
- intercalante, 91, 185, 225, 258, 355
- interfase, 97
 y transcripción, 250
- interferón, 316
- intergénico, 109
- intragénico, 109
- intrón, 277
 concepto, 77, 109, 245, 276
 disposición en el gen, 245
 eliminación por autocatálisis, 77
 en forma de lazo, 281
 en pre-tRNA, 284
 magnitud, 280
 número, 280
 polimorfismo, 368
 proporción en el genoma, 108
 tipos I y II, 77, 282
 tipos III y IV, 282
- invasividad, 414
- inversión
 cromosómica, 407
 paracéntrica, 407
 pericéntrica, 407
- IRE, 319
- IRE-BP, 319
- IRP, 319
- islote CpG, 167, 201
 en la búsqueda de genes, 380
 metilación, 272
 mutación causante de hemofilia B, 348
- isobutirilguanina, 179
- isocoro, 90
- isocromosoma, 410

isoenzima, 367
 isoesquímico, 202
 isoforma, 367
 isoleucil-tRNA sintetasa
 corrección de errores, 306
 isopeptidasa, 342
 isótopos radiactivos, 4, 180, 185,
 232
 estudio del código genético, 290
 estudio de la traducción, 300
 isozima, 367
 IVS. Véase intrón

J

JOE, 238
 jun (gen), 417

K

kanamicina, 317
 Kaposi. Véase xerodermia
 Khorana, 178
 KIP, 426
 kirromicina, 316
 Klenow (fragmento de).
 Véase DNA polimerasa I
 Klinefelter (síndrome de), 406
 Kozak (secuencia de), 308
 Kpn, 115

L

L. Véase número de enlace
 L. Véase hebra L
 L1. Véase LINE
 L-19 IVS, 77
 lactama-lactima.
 Véase tautomeria
 lactato deshidrogenasa, 437
 lacZ, 209, 214
 lamivudina, 23
 LA-PCR, 193
 lazo. Véase intrón en forma de
 LCR. Véase región de control
 del locus
 leucemia, 436
 leucina. Véase cremallera
 levógiro, 45
 Z-DNA, 55
 Li-Fraumeni (síndrome), 431
 ligamiento
 en la búsqueda de genes,
 380
 ligamiento génico, 220
 ligasa, 202
 en la reparación, 359, 361
 en la síntesis de telómeros,
 160
 mecanismo de reacción, 156
 para formar el rDNA, 207, 208
 LINE, 115, 368
 linfoma, 436
 lipofeción, 212
 liposoma, 212

líquido
 amniótico, 120
 articular, 121
 ascítico, 121
 cefalorraquídeo, 121
 pericárdico, 121
 peritoneal, 121
 pleural, 121
 serosos, 121
 sinovial, 121
 lisis celular, 132
 lisosoma, 323
 degradación de proteínas, 341
 primario, 341
 Lk. Véase número de enlace
 lobucavir, 13
 loci. *Plural de locus*
 locus, 95
 distancia entre loci, 219
 locus PI. Véase α 1-antitripsina
 Los Ángeles. Véase galactosemia
 L-PCR, 193
 luciferasa, 186
 luciferina, 186
 lumen. Véase retículo
 endoplásmico
 y véase Golgi
 luminol, 186
 luz. Véase radiación
 Lymphoprep, 124

M

maduración
 de proteínas, 326
 de mRNA, 277
 de rRNA, 282
 de rRNA y tRNA, 77
 de tRNA, 282
 postranscripcional, 244, 245,
 250, 275
 postranscripcional en la
 mitocondria, 284
 postranscripcional, diferencias
 entre procariotas y
 eucariotas, 276
 postranscripcional, regulación,
 285
 magnético. Véase microesferas
 y véase afinidad
 manosa, 330
 MAP quinasas, 422
 mapa
 de ligamiento.
 Véase mapa genético
 de restricción, 226
 de SNP. Véase polimorfismo SNP
 diferencias entre genético
 y físico, 218
 genómico, 218
 mapa físico, 218
 de alta resolución, 225
 de baja resolución, 221
 de cóntigos, 229
 de EST, 228

mapa físico
 de restricción, 226
 de STS, 228
 empleo de Southern, 227
 interpretación, 223, 227
 resolución, 226
 mapa genético, 218
 aplicación para localizar un gen,
 221
 obtención, 220
 marcador, 180.
 Véase también mapa
 EST. Véase EST
 genético, 218
 génico, 218, 228
 no génico, 218, 228
 no radiactivo, 186
 polimórfico, 221
 STS. Véase STS
 tipos, 185
 tumoral, 434, 435
 VNTR, 384
 marcaje. Véase también isótopos
 y véase fluorocromos
 con cuatro fluorocromos, 237
 con DNA polimerasa, 182
 con un indicador, 185
 con polinucleótido quinasa, 184
 de NTP en fosfato α , 182, 183,
 184
 de NTP en fosfato γ , 184, 232
 detección, 186
 directo, 181
 en cultivo, 180
 in vitro, 181
 indirecto, 185
 no nucleotídico de sondas, 185
 para la proteólisis, 341
 por cebado al azar, 183
 por traslado de la mella, 182
 terminal, 183, 232
 terminal por relleno, 184
 marca de secuencia expresada.
 Véase EST
 marco abierto de lectura, 294
 en la búsqueda de genes, 380
 marco de lectura, 288, 291, 292, 350
 desplazamiento del, 289
 Marfan (síndrome de), 395
 material genético
 extranuclear, 5
 nuclear, 5
 matriz extracelular, 122, 413, 414
 matriz nuclear.
 Véase esqueleto nuclear
 Max (proteína), 62
 Maxam y Gilbert, 231
 May-Grünwald-Giemsa (tinción
 de), 87
 β -mc, 436
 MCS. Véase sitio de clonado
 múltiple
 mdm-2 (gen), 417
 MeCP, 272
 medicina predictiva, 387

- médula ósea
muestras, 120
punción, 120
- meiosis, 97, 100, 104, 105
comparación con mitosis, 104
no-disyunción, 404, 405
- MEK, 422
- mella, 156, 202. Véase también
traslado de la mella
- Mendel
primera ley, 104
segunda ley, 105
- Meselson y Stahl, 143
- metabolismo oxidativo, 354, 377
- metafase, 99
- metaloproteasa de matriz
extracelular, 414
- metástasis, 412, 413, 414
- metilación
como señal para la reparación,
362
de aminoácidos, 328
de bases nitrogenadas, 355
de la caperuza 5', 278
de mantenimiento, 272
de novo, 272
del DNA, 271, 272
en el desarrollo embrionario,
272
regulación de la expresión génica,
272
y cáncer, 273
- metilación-restricción, 200
- metiladenina-DNA glicosilasa, 362
- N-metil-4-aminoazobenceno, 355
- N-metil-4-antraceno, 355
- metilasas, 278
asociadas a las enzimas de
restricción, 200
- O⁶-metilguanina
reparación de, 358
- metil-lisina, 328
- metilnitrosourea, 355
- metiltransferasas, 328
- metionil-tRNA iniciador, 308
- metionil-tRNA formiltransferasa,
308
- metionina, incorporación interna y
N-terminal, 308
- métodos de introducción del rDNA,
211, 213
- metotrexato, 223
- metrizamida, 124
- MFP, 424
- microesferas magnéticas, 128, 135
- β_2 -microglobulina, 436
- microinyección, 212
- micromatriz de DNA.
Véase biochip
- microsatélite.
Véase DNA microsatélite
- microscopía, 129, 186
de fluorescencia, 91, 172
- migración de células tumorales,
413, 414
- minisatélite.
Véase DNA minisatélite
- miristato, 334
- miristilación de proteínas, 335
- mitocondria, 5, 8, 323
división, 161
expresión génica, 244
generación de mutágenos, 354
rotura en la apoptosis, 432
- mitógenos, 420
- mitosis, 97, 98, 100
comparación con meiosis, 104
regulación, 423
señales de inicio, 428
sin citocinesis, 404
- mixoploidía, 403, 406
- MMP. Véase metaloproteasas de
matriz extracelular
- modelo del holoenzima para el
inicio de la transcripción,
253, 267
- modelo escalonado para el inicio
de la transcripción, 253,
267, 268
- modificación
cotraduccional, 326
de aminoácidos en proteínas,
326, 336
postraduccional, 322
- módulos de control. Véase
promotores distales
- molde, 145, 243, 247, 248
- molécula señal, 420
- mongolismo, 406
- monocistónico, 301
- monofocal, 145
- monomorfo, 368
- monosomía, 403, 404
mecanismo, 405
- mórula, 119
- mosaico, 345, 406
cromosómico, 404
- mostaza, 355
- motivo estructural
en proteínas que se unen al
DNA, 61, 265
- mRNA, 70, 71
bloqueo, 319
como muestra para clonación,
205
degradación, 286
estabilidad, 286, 319
extracción por afinidad, 139
maduración, 277
polimerasa que lo transcribe, 251
precursores de, 251
protección frente a la
degradación, 319
transporte, regulación, 318
- mRNA(guanina-N7-)
metiltransferasa, 278
- mRNA(nucleósido-2'-O)
metiltransferasa, 278
- mtDNA. Véase DNA mitocondrial
- mTERF, 257
- MTS (proteínas), 426
- mtTFA, 257
- mucoviscidosis.
Véase fibrosis quística
- muerte celular programada.
Véase apoptosis
- muesca. Véase mella
- muestras
bucales, 120
disociación, 122
tipos, 119
- multifocal, 145
- mupirocina, 316
- mutación, 142, 344
a gran escala, 346
a pequeña escala, 346
asintomática. Véase silenciosa
clasificación, 344
como origen del cáncer, 346
con cambio, desfase o
desplazamiento de
marco, 350
con sentido. Véase silenciosa
con terminación prematura.
Véase mutación sin
sentido
con terminación retrasada, 352
conservadora, 368
de sentido alterado, equivocado
o falso, 349
deleción, 347, 348
efecto beneficioso o perjudicial,
349
en DNA mitocondrial, 377
endógena, 352
espontánea, 353
exógena, 354
frecuencia, 346
génica, 344
germinal, 344
heredable, 344
incidencia, 346
inducida. Véase mutación exógena
inducida por agentes físicos, 356
inserción, 347, 348
magnitud, 346
materia prima para cambios
evolutivos, 345
neutral. Véase silenciosa
no heredable, 345
no silenciosa, 349
no sinónima. Véase de sentido
alterado
por desaminación de bases,
353
por radiación, 356
puntual, 346
silenciosa, 348, 368
sin cambio de marco, 351
sin sentido, 351
sinónima. Véase silenciosa
somática, 345
sustitución, 347
sustitución conservadora,
350

mutación
 sustitución no conservadora, 350
 transición, 347
 transversión, 347
 y degeneración del código, 295
 y metabolismo, 354
 y polimorfismo, 366
 y sustancias ambientales, 354
 mutágeno
 eliminación, 357
 endógeno, 354
 exógeno, 354
 Myc (gen y proteína), 417, 430

N

NAD⁺, 30
 cosustrato de la ligasa, 156
 NADP⁺, 30
 naftol, 186
 nalidixico (ácido), 259
 NAO, 369
 naranja de acridina, 126, 130
 NBT, 186
 necrosis, 431
 neomicina, 317
 neoplasia. Véase cáncer
 NER. Véase escisión de nucleótidos
 neurofibromatosis de tipo 1, 418
 herencia, 395
 mutación causante, 352
 neurofibromina, 395, 418
 mutada, 352
 neuronas, 97
 nf1 (gen), 269, 352
 como oncosupresor, 352, 418
 nitrato, 355
 nitrito, 355
 nitrocelulosa, 170
 nitrógeno, isótopos ¹⁴N y ¹⁵N, 143
 nitrógeno mostaza, 355
 nitrosamina, 355
 nitroso (ácido), 355
 nitrosoguanidina, 355
 no-disyunción, 404, 405
 NOR. Véase región organizadora del nucleolo
 Northern, 171
 comparado con Southern y Western, 173
 novobiocina, 259
 NPP, 186
 N-terminal. Véase proteína. extremo nucleasa
 aplicación al mapeo, 226
 aspectos generales, 198
 de restricción. Véase enzima de restricción
 en la apoptosis, 432
 S1, 169
 nucleína, 2
 núcleo celular, 2, 5
 núcleo de la DNAPol III, 149
 núcleo hierro-azufre, 319
 nucleofílico, 146, 156, 247, 313
 nucleóide, 5
 nucleolo, 251
 nucleoproteido, 31, 79
 nucleósido, 19
 absorción ultravioleta, 28, 166
 antirretroviral, 23
 componentes, 19
 -difosfato, 24
 hidrólisis química, 28
 infrecuente, 21, 72, 76, 283
 -monofosfato, 24
 propiedades fisicoquímicas, 27
 que forma ácidos nucleicos, 20
 tipos, 20
 -trifosfato, 24
 nucleosoma, 84, 85, 86
 desensamblado, 263
 liberación, 263
 posición y control de la expresión, 263
 5'-nucleotidasa
 como marcador tumoral, 437
 nucleótido
 absorción ultravioleta, 28, 166
 cíclico, 29
 complejo, 29
 complejo con Mg²⁺, 27
 componentes, 23
 didesoxi-. Véase didesoxinucleótidos
 hidrólisis química, 28
 nomenclatura, 25
 posición del enlace fosfato, 23
 propiedades fisicoquímicas, 27
 propiedades iónicas, 27
 sencillo, 24
 tipos, 23
 unido a azúcar, 332
 nuliploidia, 403
 nulismía, 403, 404
 numeración
 de los nucleótidos en genes y en RNA, 249
 número de enlace, 68
 modificación por topoisomerasas, 152
 número *n*, 93. Véanse también haploidia, diploidia, poliploidia
 variación en la fecundación, 106
 y ciclo celular, 100
 Nycoprep, 124
 nylon, 170

O

octámero de histonas, 86
 octapéptido, 251
 Okazaki (fragmentos de), 148, 153
 maduración, 156
 progreso de la síntesis, 154
 unión, 156
 oleato, 334
 oligo(dT), 139, 206, 280

oligonucleótido, 183
 empleo como sonda, 177
 empleo en PCR, 189
 síntesis química, 178
 oligosacárido, 328
 estructura química, 330
 precursor de *N*-glicanos, 333
 precursor del grupo ABO, 372
 unido por N, 330
 unido por O, 330
 olivomicina, 91
 oncogén
 comparación con oncosupresor, 419
 concepto, 416-417
 expresión dominante, 417
 que estimula el ciclo celular, 429
 oncoproteína, 416
 oncosupresor
 comparación con oncogén, 419
 concepto, 418, 419
 expresión recesiva, 419
 tipos, 418
 oocito, 103
 oogonio, 103
 operón, 252, 301
 ORF. Véase marco abierto de lectura
 orgánulo, 5, 322
 ori^c, 151
 origen
 complejo de reconocimiento, 151
 de replicación. Véase replicación, origen
 origen de transcripción
 como referencia de numeración, 249
 orina, como muestra, 122
 oro, 186, 212
 óvulo, 103
 2-oxoglutarato, 328

P

p105-Rb (proteína). Véase Rb
 p21 (proteína), 431
 p21Ras. Véase Ras (proteína)
 p53 (gen y proteína), 431
 actuación en el ciclo celular, 429
 factor de transcripción, 431
 inactivación por proteínas víricas, 431
 incidencia en el cáncer, 431
 papel dual en ciclo celular y apoptosis, 434
 PAC, 209
 pactamicina, 316
 palíndromo, 57
 como diana de restricción, 201
 en hebra sencilla, 59
 en RNA, 255
 y terminación de la transcripción, 255
 palmitato, 334
 PAP, 436

- papilomavirus, 431
- parentesco, grados de, 393
- par de bases. Véase también complementariedad
- geometría, 43
 - modelo molecular, 43
 - pares A-T, A-U y C-G, 41
- PARP, 432
- partícula de reconocimiento de la señal, 324
- partícula núcleo, 86
- Patau (síndrome de), 406
- paternidad
- pruebas de, 385
- patología molecular
- concepto, 390
- pauta de lectura. Véase marco
- PCNA, 149, 154. Véase también DNA polimerasas δ y ϵ
- secuestro por GADD, 431
- PCR, 188
- anidada, 193
 - aplicaciones, 188
 - asimétrica, 195
 - ciclo, 190
 - con adaptadores, 195
 - con comienzo en caliente, 193
 - con transcriptasa inversa, 195
 - duración, 193
 - especificidad, 193
 - etapas, 189
 - fidelidad, 193
 - fundamento, 188
 - grado de amplificación, 192, 193
 - inversa, 194
 - LA-PCR, 193
 - larga, 193
 - límite de detección, 193
 - L-PCR, 193
 - número de ciclos, 192
 - número de copias conseguidas, 192, 193
 - objetivos, 188
 - para la detección de polimorfismos, 384
 - para obtención de sondas, 176
 - para obtener DNA de hebra sencilla, 195
- reactivos, 189
- rendimiento, 192
- RT-PCR, 195
- PDGF. Véase factor de crecimiento derivado de plaquetas
- PEG. Véase polietilenglicol
- penciclovir, 13
- peptidasa, 342
- de la señal. Véase proteasa del péptido señal
- peptidiltransferasa, 78, 312, 313
- centro activo, 81
- peptidil-tRNA
- hidrólisis, 315
- péptido C, 337
- péptido señal, 323, 324, 336
- eliminación, 325
 - partícula de reconocimiento, 324
- peptidoglicanos, 329
- periodicidad
- en la estructura del DNA, 37
- permanganato potásico
- para secuenciación, 232
- peroxidasa, 185, 186
- peroxisomas, 323
- PEST, 342
- PFGE. Véase electroforesis en campo pulsante
- PGH. Véase Proyecto Genoma Humano
- pH
- efecto sobre el apareamiento de bases, 42
 - efecto sobre la desnaturalización, 164
- pintado de cromosomas, 172
- pinza deslizante, 149
- piperidina para secuenciación, 232
- pirimidina, 12
- pirofosfatasa, 146, 247
- pirofosfato, 146, 247
- pistola de genes, 212
- plantilla. Véase molde
- plásmido, 5, 209
- artificial, 209
 - características, 209
 - circularización, 194, 208
 - como vector de clonación, 204
 - F, 211
 - linealización, 177
 - origen de replicación, 209
 - pSP64, 177
 - pUC19, 209
- plegamiento de proteínas, 244, 337. Véase también carabinas
- mecanismo, 337
 - deficiencia en la fibrosis quística, 351
 - enfermedades, 337
- poli(A). Véase cola de poli(A)
- poli(A) polimerasa, 280
- poli(ADP-ribosa)polimerasa, 432
- poli(T), 139, 206, 280
- poli(U), 139, 256, 280
- poliacrilamida, 138, 173
- poliadenilación.
- Véase cola de poli(A)
- polladenilato.
- Véase cola de poli(A)
- poliamina, 34
- policación, 185, 212
- poliistrónico, 301
- policonector, 207, 209
- polietilenglicol, 212, 222
- polietilenimina, 185
- polimorfismo, 95, 96, 365.
- Véase también marcadores polimórficos
 - aplicaciones, 384, 386
 - con efecto fenotípico, 369
- polimorfismo
- concepto, 365
 - de mini- y microsatélites, 385
 - de oligosacáridos, 370
 - de un solo nucleótido. Véase polimorfismo SNP
 - detección, 381
 - en el número de repeticiones en tándem. Véase VNTR
 - en genoma mitocondrial, 377
 - en intrones, 368
 - en la longitud de los fragmentos de restricción. Véase RFLP
 - fisiológico y patológico, 366
 - genético, 365
 - genético no proteico, 349
 - génico, 365, 368, 369
 - génico y proteico, 349
 - génico no codificante, 376
 - grado de, 365
 - mecanismos de aparición, 366
 - no génico, 376
 - proteico, 349, 369, 370
 - RFLP. Véase RFLP
 - sin efecto fenotípico, 368
 - SNP, 365, 384
 - VNTR. Véase VNTR y mutación, 366
- polimorfo, 368
- polinucleótido fosforilasa, 246, 290
- polinucleótido quinasa, 184, 232
- polinucleótido para estudiar el código genético, 289
- polioma, 67
- polipéptido naciente, 81
- poliploidía, 94, 403
- polirribonucleótido, 246
- polirribosoma, 79
- polisoma. Véase polirribosoma
- polisomía, 403
- Polymorphoprep, 124
- poro nuclear, 275, 286, 318, 324
- postranscripcional
- control de la expresión génica, 284
- potenciador, 266
- pRb (proteína). Véase Rb
- precipitación
- salina diferencial, 135
- preilación de aminoácidos
- en proteínas, 335
- preparación de muestras, 118
- preproinsulina, 337
- péptido señal, 324
- pre-pro-proteína, 324, 336
- pre-proteína, 324
- prequimotripsinógeno, 336
- pre-RNA, 244, 245, 250, 275
- con varios productos génicos, 276
- de transferencia, 284
- maduración. Véase maduración postranscripcional
- mitocondrial, 257, 284
- mixto para tRNA y rRNA, 283

- pre-RNA
 productos génicos, 251
 ribosómico, 76, 80, 251, 283
 ribosómico de *Tetrahymena*, 77
 pretraduccional, control de
 la expresión génica, 284
 pretranscripcional, control de
 la expresión génica, 263
 pretripsinógeno, 336
 Pribnow (caja de). Véase TATA
 primasa, 147, 148, 152.
 Véase también DNA
 polimerasa α
 primosoma, 152
 prión, 337
 probando, 393
 procarionte. Véase procarionta
 procarionta
 diferencias con eucariota, 5
 genoma, 5
 procesamiento postranscripcional.
 Véase maduración
 postranscripcional
 procesamiento proteolítico de
 proteínas.
 Véase proteólisis
 procesividad, 148, 149
 procolágeno, 328
 producto génico, 245
 nomenclatura, 246
 profase, 99
 proflavina, 355
 prohormona convertasas
 PC2 y PC3, 337
 proinsulina, 337
 proliferación, 213
 control por señales
 extracelulares, 420
 control negativo, 418
 control positivo, 416
 equilibrio con muerte celular,
 414
 monoclonal, 413
 prolil hidroxilasa, 328
 prolina
 hidroxilación, 328
 prometáfase, 99
 promotor, 245, 252, 253, 264
 basal, 265
 basal, complejo plurimolecular, 267
 basal, reconocimiento por TBP,
 267
 de clases I, II y III, 265
 desnaturalización, 253
 distal, 266
 en vectores de expresión, 215
 liberación del, 254
 metilación, 272
 mitocondrial, 257
 potenciador, 266
 proximal, 265
 silenciador, 266
 pronúcleo, 106
 propagación celular, 213
 propidio. Véase yoduro de propidio
 propositus, 393
 pro-proteína, 324, 336
 protamina, 34, 84
 proteasa, 122, 342, 414
 del péptido señal, 324
 cisteinproteasa, 432
 digestiva, 336
 en la apoptosis, 432
 inhibición, 376, 414
 lisosomal, 341
 proteasoma, 342
 protección
 de grupos funcionales, 178, 179
 proteína, 2
 activación proteolítica.
 Véase proteólisis
 como producto génico.
 Véase producto génico
 de adhesión, 413, 437
 de anclaje, 325
 de choque térmico, 339
 de membrana, 324
 de pausa, 256
 de replicación A. Véase RPA
 de secreción, 324, 336
 de unión a hebra sencilla, 150,
 151
 de unión a Inr, 270
 de unión a la caperuza.
 Véase CBP
 de unión a TATA, 267, 271
 degradación, 340, 341
 detectora, 181, 185
 diferencias entre tipos celulares,
 262
 extremos amino o N-terminal
 y carboxilo o C-terminal,
 300
 funcional, 322
 G, 421
 glicosilación, 328
 interacción con DNA, 60
 ligante, 181
 ligante de IRE, 319
 ligante del telómero, 92
 naciente, 300, 314, 338
 no histonas, 84
 plegamiento.
 Véase plegamiento de proteínas
 priónica, 337
 receptora de AMP cíclico, 62
 receptora de SRP, 325
 receptora del ribosoma, 325
 reguladora del hierro, 319
 síntesis de. Véase traducción
 supresoras de tumores.
 Véase oncosupresores
 terminadoras de la transcripción,
 256
 transportadora de cloruro, 351
 Véase regulador de la
 conductancia
 transmembrana
 ubiquitinadas, 341
 vida media, 340
 proteína-disulfuro isomerasa, 335,
 338
 proteína fosfatasa, 327
 proteína quinasa, 327, 425
 dependiente de ciclina.
 Véase Cdk
 TFIIH, 267
 proteoglicanos, 329
 proteólisis, 336
 de la DNAPI, 147
 de proteínas marcadas con
 ubiquitina, 341-342
 Véase también proteasa
 proteoma, 217
 proteómica, 217
 protooncogén
 concepto, 416, 417
 tipos, 416
 proximal. Véase promotor
 Proyecto Genoma Humano, 217
 características, 218
 etapas, 239
 proyectos privados, 240
 PrP, 337
 pseudogén, 112
 y transcripción, 250
 pseudomónico (ácido), 316
 pseudouridina, 21, 72
 psoraleno, 359
 pSP64. Véase plásmido
 PTE, 255
 pUC19. Véase plásmido
 puente
 de hidrógeno. Véase enlace
 disulfuro. Véase enlace
 punto de restricción, 423
 punto R, 423
 purina, 12
 puromicina, 21, 316, 317
- ## Q
- queratinas, 436
 quiescencia, 98
 quimera, 406
 quimioluminiscencia, 186
 quimotripsina, 336
 inhibición por antitripsina, 376
 quimotripsinógeno, 336
 quinacrina, 91
 quinasa. Véase también
 proteína quinasa
 activadora de Cdk, 427
 inhibidora de Cdk, 427
- ## R
- Rab (proteína), 421
 radiación
 ionizante, 356
 rayos gamma, 356
 rayos X, 356
 ultravioleta, 356
 visible y reparación del DNA,
 357

- radicales libres. Véase especies reactivas de oxígeno
- Raf (proteína), 422
- ramificación, centro de, 280
- Ras (gen y proteína), 417
 en neurofibromatosis tipo 1, 352
 farnesilación de, 335
 función, 421
 mutada, 422
- rasgo falciforme, 396
- rasgos, 95
- razón de asimetría, 39
- Rib (gen y proteína), 430
 actuación sobre el ciclo celular, 429
 como sustrato de Cdk, 430
 fosforilación, 430
 inactivación por proteínas víricas, 431
 inactivación y cáncer, 430
- rDNA, 203, 204
 formación, 207
 introducción en la célula anfitriona, 211, 213
 subproductos, 208
- reasociación del DNA. Véase Cot y véase renaturalización
- recambio proteico, 340
- receptor
 de factores de crecimiento, 417
 de factores inhibidores del crecimiento, 418
 de hormonas tiroideas, 417
 de transferrina, 319
- recesivo, 95, 394
- Recklinghausen (enfermedad de). Véase neurofibromatosis
- recombinación, 106, 142
 fracción. Véase frecuencia
 frecuencia. Véase frecuencia
 homóloga, 106, 218
 meiótica, 218
 no homóloga o desigual, 367
 relación con el mapeo, 218
 y segregación de alelos, 218
- recuento de células viables, 129
- recuperación de bases, vía metabólica, 223
- reemplazamiento, método de preparación de virus recombinantes, 210
- región
 de control del locus, 111
 estructural.
 Véase gen, secuencias reguladora.
 Véase gen, secuencias organizadora del nucleolo, 91
- regiones hipervariables en el mtDNA, 377
- regiones internas de control, 271
- regulación
 por fosforilación reversible, 425
 pretraducción, 319
 traducción, 318
- regulador de la conductancia transmembrana, 396
- reguladora (secuencia génica). Véase gen, secuencias
- relación de codificación, 288
- relleno, 184
- renaturalización, 163, 167
- rendimiento de la PCR, 192
- reparación, 142, 356
 y actividad 5'-exonucleasa, 147
 de apareamientos incorrectos, 362
 de bases metiladas, 357
 del uracilo en el DNA, 362
 por escisión, esquema global, 358
 por recombinación, 362
 posreplicación, 362
 sistemas enzimáticos, 356
 tipos, 356
 Véase también escisión
- repetición. Véase DNA repetitivo
 directa, 58
 especular, 58
 invertida. Véase palíndromo
 terminal. Véase telómero
- replicación, 2, 98, 141
 autónoma, 206
 bidireccional, 144
 bifocal, 161
 bloqueo por GADD o p21, 431
 características comparadas con transcripción, 248
 ciclo celular, 142
 cofactores, 145
 comparación entre procariotas, eucariotas y mitocondrias, 161
 complejo de. Véase replisoma
 conservadora, 142, 144
 del RNA, 2, 147
 demostración del mecanismo, 143
 dirección, 146
 dispersante, 142
 elongación. Véase elongación de la replicación
 en la PCR, 189
 enzimología, 145
 errores, 352, 362
 final. Véase terminación
 hipótesis, 142, 144
 horquilla, 145, 150, 151
 inicio, 150
 mecanismo de la reacción, 146
 mitocondrial, 160
 monofocal, 145
 multifocal, 145
 no simultánea, 161
 origen, 92, 145, 150, 211
 origen mitocondrial, 160
 origen plasmídico, 209
 orígenes múltiples en un cromosoma, 157
 reacción comparada con transcripción, 247
- replicación
 reacción global, 146
 relación con el ciclo celular, 100
 secuencial, 144
 semiconservadora, 142, 144
 semidiscontinua, 153
 señales de inicio, 428
 simultánea, 144
 sustratos, 145
 terminación. Véase terminación
 unidad de. Véase replicón
 unidireccional, 161
 velocidad, 149
- replicón, 145, 157
- replisoma, 151, 152
 modelo, 155
- resistencia a antibióticos, 206, 209, 214, 316
- restricción. Véase enzimas de restricción
- restricción-modificación, 200
- restringidas. Véase enzimas de restricción
- retardada (hebra), 152-154
- retículo endoplásmico, 322
 lumen, 324
- reticulocitos, para estudiar la traducción, 300
- retinoblastoma. Véase Rb
- retrasada (hebra). Véase retardada
- retrotransposición, 368
- retrotransposón, 368
- retrovirus, 2, 210, 244
- revelado del DNA tras electroforesis, 138
- reversión directa de la mutación, 357
- RFC, 154. Véase también DNA polimerasas δ y ϵ
- RFLP, 381
- Rho (proteína), 421
- riboforina, 325
- ribonucleasa
 como marcador tumoral, 437
 como proteína modelo de plegamiento, 338
 Véase también RNasa
- ribonucleoproteínas nucleares pequeñas. Véase snRNP
- ribosa, 11
- ribosoma, 79
 centros activos, 81
 composición, 80, 81
 deslizamiento en el inicio de la traducción, 309, 310
 deslizamiento en la elongación, 313
 disociación durante la traducción, 309, 315
 en mitocondrias y cloroplastos, 79
 interacción con mRNA, 81
 separación y reasociación de subunidades, 309, 310
 sitios A, P y E, 311

ribosoma
 subunidades, 80, 81, 308, 309
 tamaño, 79, 80
 ribosomas.
 Véase sondas de RNA
 ribotimidina, 72
 ribozima, 76
 auto-ajuste, 282
 RNA autocatalítico, 282
 transpeptidasa, 313
 Rich-Dickerson, 52
 ricina, 316
 rifamicinas, 258
 rifampicina, 251, 258
 rigor, 168
 RNA, 2
 7SL, 114, 324
 abundancia, 70
 catalítico. Véase ribozimas
 citoplásmico pequeño.
 Véase scRNA
 como producto génico.
 Véase producto génico
 composición de bases, 69
 cromosómico, 2
 de transferencia. Véase tRNA
 diversidad de estructuras, 70
 en orgánulos, 70
 estabilidad, 280, 286
 estructura, 69
 estructura secundaria, 70, 73
 funcional. Véase RNA maduro
 genómico, 5
 maduración. Véase maduración
 postranscripcional
 maduro, 245, 251, 275
 mensajero. Véase mRNA
 mitocondrial, 257
 mitocondrial, maduración, 284
 nuclear heterogéneo.
 Véase hnRNA
 nuclear pequeño. Véase snRNA
 naciente, 247, 254
 poliadenilado.
 Véase cola de poli(A)
 precursor. Véase pre-RNA
 ribosómico. Véase rRNA
 secuenciación, 238
 tamaño, 70
 tipos, 69, 70
 transporte, 244, 280
 ubicación en la célula, 70
 vida media, 280
 RNA polimerasa, 84, 244
 activación por el complejo
 de inicio, 267
 cloroplástica, 251
 en la transposición, 368
 especialización, 251, 265
 inhibición, 251
 localización subcelular, 251
 mitocondrial, 251, 257
 pausas, 256
 procariótica, 250
 procariótica, enzima mínima o
 núcleo, 251, 254

RNA polimerasa
 procariótica, holoenzima, 254
 procariótica, subunidad σ , 251,
 254
 SP6, 177
 subunidades, 251
 tipos, 250
 transcritos primarios
 sintetizados, 251
 I, 251, 271
 II, 251, 268
 II, dominio C-terminal.
 Véase CTD
 III, 251, 271
 RNA replicasas, 2, 147
 RNAPol. Véase RNA polimerasas
 RNasa, 199
 como contaminación, 171
 D, 283
 de *Bacillus cereus*, 238
 degradación del mRNA, 286
 H1, 156
 H1, síntesis de telómeros, 160
 P, 77, 283
 Phy M, 238
 T1, 238
 U2, 238
 I, 199, 283
 III, 283
 ro (proteína), 255
 análogos eucarióticos, 257
 Robertson. Véase translocación
 rodamina, 238
 rodopsina, 334
 rojo neutro, 130
 romo. Véase extremo romo
 ROS. Véase especie reactiva de
 oxígeno
 ROX, 238
 RPA, 151, 360
 rRNA, 70, 76
 5S, 70, 76, 80
 5.8S, 70, 76, 80
 16S, 70, 76, 80
 18S, 70, 76, 80, 91
 23S, 70, 76, 78, 80
 28S, 70, 76, 78, 80, 91
 28S, como transpeptidasa, 313
 generación por maduración, 283
 genes de, 111
 maduración, 282
 polimerasas que lo transcriben,
 251
 precursores de, 251
 RT-PCR, 195

S

saco
 amniótico, 120
 vitelino, 119
 salientes 5' y 3' en los extremos de
 restricción, 201
 saliva, muestras, 120
 Sanger, 235

sangre
 manchas, 120
 fetal, 120
 periférica, 120
 sarcosina, 259
 satélite. Véase DNA satélite
 satélite cromosómico.
 Véase cromosoma
 SCC, 436
 scRNA, 70
 secuenciación, 235
 secuencia
 cis. Véase cis
 cos, 210
 Inr, 253
 TATA, 253
 de etiquetado.
 Véase péptido señal
 de replicación autónoma, 211
 estructural.
 Véase gen, secuencia
 génica. Véase gen, secuencia
 intercalada. Véase intrón
 líder. Véase péptido señal
 promotora. Véase promotor
 reguladora.
 Véase gen, secuencias
 señal. Véase péptido señal
 subtelomérica, 92
 telomérica. Véase telómero
 trans. Véase trans
 secuenciación, 218, 231
 de genomas, 241
 de RNA, 238
 didesoxi. Véase secuenciación,
 método enzimático
 en la búsqueda de genes, 380
 enzimática, variantes, 238
 método enzimático, 235
 método químico, 231
 por terminación de cadena.
 Véase secuenciación,
 método enzimático
 shof-gun, 240
 secuencial. Véase replicación
 sedimentación, 122
 segregación
 de alelos, 218
 de cromosomas y cromátidas,
 104, 105, 218
 segundo código genético, 304
 selección
 de células híbridas, 223
 de clones recombinantes, 213
 fenotípica, 214
 negativa, 129
 positiva, 129
 semen, 120
 semiconservadora.
 Véase replicación
 semidiscontinua.
 Véase replicación
 señalización, tipos, 420
 separación de células, 122
 serina/teonina quinasa, 327, 424, 425

Shine-Dalgarno (secuencia de), 308
 showdomicina, 316
 siálico (ácido), 330
 sigma. Véase RNA polimerasa procariótica
 silenciadores, 266
 simultánea. Véase replicación
 SINE, 114
 síntesis de novo de nucleótidos, 223
 síntesis fuera de programa, 358
 sis (gen), 416
 sistema endomembranoso, 322
 sitio
 A o aminoacilico, 311, 313
 de clonado múltiple, 207, 209
 de restricción, 199
 de restricción, ejemplo, 209
 de secuencia marcada.
 Véase STS
 E o de salida, 311, 313
 marcados únicos. Véase STS
 P o peptídico, 311, 313
 SL1, 271
 slot-blot en hibridación, 170
 snip. Véase polimorfismo SNP
 SNP. Véase polimorfismo SNP
 snRNA, 70, 251, 271, 280
 tipos, tamaño y función, 282
 snRNP, 280
 tipos, tamaño y función, 282
 sobrecruzamiento.
 Véase recombinación
 solapamiento de codones, 288
 solenoide. Véase fibra de 30 nm
 sonda, 168
 clasificación, 175
 de DNA, 176
 de RNA, 177
 diseño, 179
 fluorescentes, 172, 224
 marcage, 180
 obtención por clonación, 176
 obtención por PCR, 176
 para detectar la anemia
 falciforme, 383
 preparación, 175
 síntesis, 183, 184
 sintéticas, 177
 soporte sólido, unión a, 178
 Southern, 171
 aplicación al mapeo físico, 227
 comparado con Northern y
 Western, 173
 interpretación de resultados, 172
 para detectar la anemia
 falciforme, 383
 para detectar RFLP, 381, 383
 para detectar VNTR, 383
 SP1, 269
 SP6, 177
 SREBP, 432
 SRP, 324
 SRP-R, 325

SSB. Véase proteínas de unión a
 hebra sencilla
 STS en mapa físico, 228
 subclonación, 215
 subtilisina, proteólisis de la
 DNApol I, 147
 suicidio celular. Véase apoptosis
 superenrollamiento, 65, 83, 87
 aspectos teóricos, 68
 demostración experimental, 67
 negativo, 67, 152
 positivo, 67
 superhélice, 66
 superóxido, 354, 357
 superóxido dismutasa, 357
 surco mayor, 46, 48
 unión de proteínas, 61
 surco menor, 46, 48
 unión de proteínas, 267
 survivinas, 433
 susceptibilidad tumoral, 413, 434
 sustancia H. Véase antígeno H
 sustitución. Véase mutación
 método de preparación de virus
 recombinantes, 210
 SV40, 210
 svedberg. Véase coeficiente de
 sedimentación
 SII, 255
 SIII, 255

T

TAF. Véase factores asociados
 a TBP
 talasemia, 396
 tallo, 70
 TAMRA, 238
 tándem, 109, 383
 TATA, 253, 265, 270
 reconocimiento por TBP, 267
 tautomería, 14
 relación con la mutación, 353
 Tay-Sachs (enfermedad de), 350, 396
 TBP. Véase proteína de unión
 a TATA. Véase proteína
 ligante del telómero
 3TC. Véase lamivudina
 Telofase, 99
 telomerasa, 147, 158
 ausencia, 160
 componentes, 158
 función, 160
 mecanismo de acción, 159
 y cáncer, 160
 y proliferación celular, 160
 telómero, 87, 91, 92
 acortamiento, 157, 158, 160
 en vectores de clonación, 211
 funciones, 92
 replicación, 157
 secuencia, 159
 síntesis, 159
 y DNA minisatélite, 114
 y DNA satélite, 113
 telómero
 y envejecimiento, 160
 temperatura
 de fusión del DNA. Véase T_m
 en la PCR, 189
 y desnaturalización, 164
 templado, 168
 en la PCR, 189
 TER. Véase telomerasa,
 componentes
 terapia antitumoral
 y telomerasa, 160
 terminación
 de la replicación, 157
 de la traducción, 314
 de la transcripción, 255
 de la transcripción mitocondrial,
 257
 dependiente de p, 255
 independiente de p, 255
 prematura de la traducción, 351
 retrasada de la traducción, 352
 terminación de cadena. Véase
 secuenciación enzimática
 terminador, 252, 255
 termociclador, 189
 termoestable, 189, 235
 terpeno, 335
 tetraciclina, 316
 gen de resistencia, 214
 tetrahidrofolato, 223, 308
 Tetrahymena, 77
 tetraploidía, 94, 403, 404
 tetrazolio, 131, 186
 T_m . Véase T_m
 TF. Véase factores de
 transcripción
 TGB, 436
 TGF- β , 418
 timidina, 20
 en medio HAT, 223
 triflato, 132
 timidina quinasa, 223
 timina, 12
 dímero, 356, 359
 timina glicol, 354
 reparación de, 360
 tinción, 87, 90, 91. Véanse
 también Giemsa y May
 del DNA, 138
 NOR, 91
 tiouridina, 72
 tiroglobulina, 436
 tirosina quinasa, 327, 425
 TK. Véase timidina quinasa
 T_m , 166
 en la PCR, 189
 y contenido en G+C, 166
 TMB, 186
 topogénesis, 322
 topoisomerasa, 66, 84, 87, 151
 de tipo I, 152
 en la transcripción, 254
 topología, 68
 toxina diftérica, 317

- TR. Véase telomerasa, componentes
- traducción, 2, 244, 245
características, 300
conexión espacial y temporal con la transcripción, 250
elongación. Véase elongación de la traducción
energética, 315
fases o etapas, 302
frecuencia, 319
inhibidores, 316
inicio. Véase inicio de la traducción
macromoléculas implicadas, 300
regulación, 318-320
resumen, 302, 315
sentido, 300
tasa de error, 307
terminación, 314
velocidad, 318
tráfico de proteínas, 322-324
- trans
elementos o secuencias, 265
factores, 265
Véase también Golgi
- transcripción, 2, 243, 245
asimetría, 249
características comparadas con replicación, 248
cofactores, 247
comparación entre procariotas y eucariotas, 250
conexión espacial y temporal con la traducción, 250
del genoma mitocondrial, 257
dirección, 247
e interfase, 250
elongación. Véase elongación de la transcripción
energética, 246
enzimología, 246
etapas, 252
hebras positiva y negativa, 248
hebras codificante y no codificante, 248
hebras con sentido y antisentido, 248
hebras informativa y no informativa, 248
hebras molde y no molde, 248
hebras transcrita y no transcrita, 248
inhibidores, 258
inicio. Véase inicio de la transcripción
mecanismo de la reacción, 247
origen. Véase origen de transcripción
promotor. Véase promotor
reacción comparada con replicación, 247
reacción global, 246
reacción inicial, 254
regulación, 261
- transcripción
regulación epigenética, 271
regulación genética, 264
regulación por señales, 266
regulación positiva y negativa, 266
regulación postranscripcional, 284
regulación y accesibilidad del DNA, 263
regulación, terminología, 265
relación con la condensación del DNA, 250
sentido contrario en cada hebra, 249
sustratos, 246
terminación, 255
terminología, 248, 249
unidad de, 252
y eucromatina, 250
y fibra de 10 nm, 250
- transcripción inversa, 2
transcriptasa.
Véase RNA polimerasa
- transcriptasa inversa, 147, 244
en la transposición, 368
inhibidores, 23
para PCR, 196
para preparar cDNA, 206
para secuenciar DNA, 235
para secuenciar RNA, 239
- transcrita. Véase hebra transcrita
transcrito primario.
Véase pre-RNA
- transducción de señal, 417, 418, 420
esquema global, 421
- transesterificación, 281, 303
- transfección, 212
- transferasa terminal, 206
- transferencia tras electroforesis, 171, 173
- transferrina, receptor de, 319
- transformación bacteriana, 211, 212
- transfosforilación, 281
- transfusión, compatibilidad, 370
- transición, 347
- translocación
de proteínas del citosol al retículo endoplásmico, 324
de Robertson, 409
del ribosoma, 312, 313
papel de las carabinas, 337
recíproca, 409
- translocasa. Véase eEF-2 y véase riboforina
- transpeptidación, 312, 313
- transpeptidasa.
Véase peptidiltransferasa
- transporte
a través de poros nucleares, 244
de mRNA, 318
regulado, 323
- transporte
transmembrana, 323
vesicular, 323
- transposasa, 367
- transposición, 367
- transposón, 367
- transversión, 347
- traslado de la meiosis, 182
- trébol, estructura del tRNA, 74
- tricotodistrofia, 363
- trifosfato, 10
- triple hélice, 60
- tripleto. Véase codón
- triploidia, 94, 403
- tripsina, 336
inhibición por antitripsina, 376
proteólisis de la DNAPI, 147
- tripsinógeno, 336
- triptófano.
biosíntesis deficiente, 214
- trifosfato, 29
- trisomía, 403, 404, 406
del cromosoma 13, 406
del cromosoma 18, 406
del cromosoma 21, 345, 406
de cromosomas sexuales, 406
mecanismo, 405
- tritilo, 179
- tRNA, 70, 71
especificidad por aminoácido, 71
estructura terciaria, 75
isoaceptor, 304
maduración, 282
modelo molecular, 75
nomenclatura, 71
número de moléculas diferentes, 297
papel adaptador en la traducción, 289
polimerasa que lo transcribe, 251
precursores de, 251
separación del polipéptido al terminar la traducción, 315
sinónimo, 304
tRNA^{met}, 308
tRNA^{met} y tRNA^{met}, 308
complejo ternario con eIF-2 y GTP, 309, 310
- tRNA nucleotidiltransferasa, 283
- trofoblasto, 119
- trombina, carboxilación, 327
- troponina T, gen y ajuste alternativo, 286
- TRP1, 214
- TRT. Véase telomerasa, componentes
- TTD-A, 363
- tumor. Véase también cáncer maligno, 414
primario benigno, 413
secundario, 413, 414
tumorigénesis, 413

tungsteno, microesferas para
clonación, 212
Turner (síndrome de), 404

U

U1, U2, U3, U4, U5, U6. Véase
snRNA y véase snRNP
UBF, 271
ubicación del aminoacil-tRNA, 312
ubiquitina, 341
ubiquitina proteína ligasa, 341
ubiquitinación, 342
UCE, 271
UDP-galactosa, 332
UDP-galactosa epimerasa, 374
UDP-glucosa, 332
UDP-glucosa pirofosforilasa, 374
UDP-N-acetilgalactosamina, 332
UDP-N-acetilglucosamina, 332
ultracentrifugación, 136
isopícnica, 67
ultravioleta. Véase radiación
unidad
de codificación, tamaño, 288
de repetición.
Véase DNA repetitivo
de replicación. Véase replicón
de transcripción, 252
de transcripción mitocondrial, 258
unidireccional. Véase replicación
universalidad.
Véase código genético
uracilo, 12
reconocimiento en el DNA, 362
uracilo-DNA glicosilasa, 362
urea, 134, 165, 169
uridilato, 27
uridililtransferasa, 374
uridina, 20

V

vall-tRNA sintetasa,
corrección de errores, 307
valor C, 6, 94
variación en la fecundación, 106
variación en la meiosis, 104

valor C
variación en el ciclo celular,
100
variabilidad genética.
Véase polimorfismo
vascularización, 414
vector
ARS/CEN/TEL, 211
BAC, 211
clasificación, 208
cósmido, 210
de clonación, 203, 206
de expresión, 206, 215
de inserción, 206, 210
de propagación, 206
de sustitución, 210
plasmídico, 209
preparación, 207
propagación, 209
tamaño del inserto admitido,
209
tipos, 208
vírico, 210, 212
YAC, 211
vellosidades coriónicas, 120
velocidad de sedimentación, 122,
124
vesículas, 323
viabilidad, 129
vida media
de proteínas, 340
del RNA, 286
VIH. Véase virus de la
inmunodeficiencia
humana
viroides, 78
virus. Véase también bacteriófagos
adenovirus, 431
cápsida, 4
de la inmunodeficiencia
humana, 23
del mosaico del tabaco, 2
del papiloma, 431
empleo como vector.
Véase vector vírico
recombinantes, 210
SV40, 431
tumores, 431

viscosidad, 34
variación en la desnaturalización
del DNA, 165
vitamina B₁₂. Véase cobalamina
vitamina C, 328
vitamina K, 327
VNTR, 221, 383

W

WAF1 (proteína), 426
Watson y Crick, 2. Véase también
enlace de hidrógeno
modelo para el DNA, 39
Western, 173
comparado con Southern y
Northern, 173

X

xantina, 13
como resultado de la
desaminación, 353
xerodermia pigmentosa o
de Kaposi, 360, 363
X-gal, 186, 214
XP. Véase xerodermia pigmentosa
XPA, 360
XPC, 360
XPF, 360
XPG, 360

Y

YAC, 209, 211
yoduro de propidio, 130, 131
YY1, 270

Z

zalcitabina, 23
Z-DNA, 55
modelos moleculares, 55
zeiosis, 432
zidovudina, 23
zig-zag
disposición en el Z-DNA, 56
zimógenos, 336
zinc. Véase dedo de zinc

